Combinatorial aspects of root lattices and words

Thesis

Combinatorial Aspects of Root Lattices and Words

Manuela Heuer
Dipl.-Math.

A thesis submitted for the degree
Doctor of Philosophy
to

Department of Mathematics and Statistics
The Open University
Milton Keynes, United Kingdom

May 2010

Abstract

This thesis is concerned with two topics that are of interest for the theory of aperiodic order. In the first part, the similar sublattices and coincidence site lattices of the root lattice A_4 are analysed by means of the quaternion algebra $\mathbb{H}(\mathbb{Q}(\sqrt{5}))$. Dirichlet series generating functions are derived, which count the number of similar sublattices, respectively coincidence site lattices, of each index.

In the second part, several strategies to derive upper and lower bounds for the entropy of certain sets of powerfree words are presented. In particular, Kolpakov's arguments [**49**] for the derivation of lower bounds for the entropy of powerfree words are generalised. For several explicit sets we derive very good upper and lower bounds for their entropy. Notably, Kolpakov's lower bounds for the entropy of ternary squarefree, binary cubefree and ternary minimally repetitive words are confirmed exactly.

Acknowledgements

First of all, I wish to thank my supervisors Uwe Grimm and Michael Baake for the opportunity to write this thesis and their support in its preparation. Further, I would like to thank Peter Zeiner for his cooperation and careful reading of various manuscripts. I would also like to thank Roman Kolpakov for a very helpful email discussion. Furthermore, I wish to thank Andrey Umerski for being my third supervisor and Paul Upton for being my third party monitor.

The excellent research environment provided by the Department of Mathematics and Statistics at the Open University is gratefully acknowledged. Moreover, I would like to thank the entire research group "Mathematik in den Naturwissenschaften" at the University of Bielefeld for providing an always welcoming atmosphere. A big thank you to Christian Huck and Svenja Glied for their understanding and encouragement not only but in particular during the time we shared in Milton Keynes. I am very grateful for their comments on the manuscript.

Finally, I wish to express my gratitude to all the people who have supported me during the last years, particularly to my friends and family. Last but not least, I would like to thank my partner David for his continuous encouragement and unfailing patience.

Publications

Parts of my research were done in collaboration with members of the research group of my external supervisor Michael Baake (University of Bielefeld, Germany) with my internal supervisor Uwe Grimm and with Robert V. Moody (University of Victoria, Canada). The following parts of my thesis have led to joint publications.

The content of Chapter 2 appeared in [**9**]. The analysis of the coincidence rotations in Chapter 3 was published in [**8**]. A summary of these results with respect to the dual of the A_4 root lattice also appeared in [**40**]. The classification of the CSLs in Chapter 3 is published in a summarised form in [**41**]. Moreover, Chapter 4 is based on [**37**].

Note

The main results of Chapter 2, which are summarised in Section 2.2.3, are also contained in my "Diplomarbeit: Ähnlichkeitsuntergitter des Wurzelgitters A_4", which was submitted to the University of Bielefeld in 2006, see [**39**]. However, in this thesis these results are derived from an alternative approach, which is also published in a joint paper [**9**] with my external supervisor Michael Baake and Robert V. Moody. Moreover, for reasons of completeness Theorem 2.6 is recalled from [**39**].

Contents

List of Figures

Introduction

Root lattices and words play an important role in the theory of aperiodic order, which is concerned with systems that display order without periodicity, see [**7**] for a general introduction to the field. This thesis is split into two parts; Part I is about a particular root lattice, the root lattice A_4, while Part II deals with powerfree words.

Consider a general lattice Γ in the Euclidean space $\mathbb{R}^d$ and let σ be a similarity of $\mathbb{R}^d$, i.e. a non-zero linear map of $\mathbb{R}^d$ with $\langle \sigma(u), \sigma(v) \rangle = c\langle u, v \rangle$ for all $u, v \in \mathbb{R}^d$, where $c > 0$ and $\langle .|. \rangle$ denotes the standard Euclidean scalar product. If $\sigma(\Gamma) \subset \Gamma$, then the sublattice $\sigma(\Gamma)$ is called a *similar sublattice (SSL)* of Γ. Every lattice Γ possesses trivial SSLs of the form $m\Gamma$, where $m \in \mathbb{N}$. Lattices with a rich point symmetry structure, like root lattices, have many non-trivial SSLs in addition. Several lattices have already been investigated with respect to their SSLs, compare [**11, 12, 23, 14**] and references given there. In [**23**], the possible indices of SSLs for many root lattices, including the root lattice A_4, were derived. However, the question how many different SSLs of each index exist, has remained open for the root lattice A_4. In Part I we answer this question by deriving a Dirichlet series generating function for the number of SSLs of each index; compare also [**9**].

The classification of SSLs of a lattice Γ is closely related to that of its coincidence site lattices; see for example [**35**]. A finite-index sublattice of Γ of the form $\Gamma \cap R\Gamma$, where R is an orthogonal map of $\mathbb{R}^d$, is called a *coincidence site lattice (CSL)* of Γ. CSLs are used in crystallography in the description and understanding of grain boundaries, compare [**3**] and references given there. For many lattices in dimension $d \leq 4$, except for the

root lattice A_4, the arithmetic function which counts the number of CSLs of each index has been derived; see for instance [**71, 3, 88, 13, 15**]. In Part I we derive this arithmetic function and the corresponding Dirichlet series for the root lattice A_4; compare also [**8, 40, 41**].

Generally, the root lattice A_4 is of particular interest, because it forms the natural setting, in the sense of a minimal embedding, for the description of the Penrose tiling as a cut and project set, see for example [**10**]. The Penrose tiling is a classical example for an aperiodic tiling. Some of the various other applications of the root lattice A_4 are described in [**23**].

Several attempts have been made to get further insight into the theory of classifying SSLs and CSLs for a general lattice or module in $\mathbb{R}^d$, see [**92, 91, 34, 42**] and references given there for recent publications. However, results for general lattices or modules remain sparse. Another generalisation is the analysis of multiple CSLs, i.e. finite-index sublattices of a lattice Γ which have the form $\Gamma \cap R_1\Gamma \cap \ldots \cap R_m\Gamma$ where $R_1, \ldots, R_m$ are isometries, see [**6, 88, 89**].

Substitution sequences often provide interesting models for aperiodic systems in one dimension. For example, the famous Thue-Morse morphism

$$\varrho\colon \begin{array}{rcl} 0 & \mapsto & 01 \\ 1 & \mapsto & 10 \end{array}, \tag{0.1}$$

which was first defined by Thue [**84, 83**] at the beginning of the 20th century and later rediscovered by Morse [**61**], generates, via iteration on the initial word 0, the infinite word

$$0110100110010110100101100110100110010110011010010110100110010110\ldots.$$

This word clearly shows some kind of order, but already Thue proved that it is cubefree [**84**], which means that it does not contain any subword of the form $0^3 = 000$, $1^3 = 111$, $(01)^3 = 010101$, $(10)^3 = 101010$ and so on. Its

cubefreeness is a consequence of the fact that the morphism ϱ maps cubefree words to cubefree words. A morphism with this property is called cubefree.

In general, the iteration of a powerfree morphism is a convenient way to produce infinite powerfree words. However, systems produced in this way have zero combinatorial entropy (see Definition 4.12). A natural generalisation to an interesting set with positive entropy is provided by the set of all powerfree words.

The investigation of powerfree words is one particular aspect of combinatorics on words. The book series [**55, 56, 57**] gives a comprehensive overview of the field. Since its initiation by Thue, it has attracted considerable interest particularly in the past decades [**16, 25, 18, 19, 44, 51, 36, 26, 76, 17**], and continues to do so, see [**75, 64, 77, 49, 65, 66**] for some recent work. Beyond the field of combinatorics on words and aperiodic order, substitution sequences, such as the Thue-Morse sequence, have been investigated for instance in the context of symbolic dynamics [**72, 33, 1**].

Part II is about the combinatorial entropy of the set of powerfree words. Due to the fact that every subword of a powerfree word is again powerfree, the entropy of powerfree words exists as a limit. It is a measure for the exponential growth rate of the number of powerfree words of length n. So far neither an explicit expression for the entropy of powerfree words nor an easy way to compute it numerically is known. Nevertheless, there are several strategies to derive upper and lower bounds for this limit. Upper bounds can be obtained, for example, by the enumeration of all powerfree words up to a certain length. Until recently, all methods to achieve lower bounds relied on powerfree morphisms. However, the lower bounds obtained in this way are not particularly good, since they are considerably smaller than the upper bounds, which are close to numerical estimates of the entropy. A completely different approach, introduced recently by Kolpakov [**49**], provides surprisingly good lower bounds for the entropy. Here, several methods to

derive upper and lower bounds including Kolpakov's are explained in detail and applied to a number of examples.

The thesis is organised as follows. Chapter 1 introduces the general notation and terminology of Part I. The A_4 root lattice is presented in a realisation ensuring that it is contained in a particular maximal order of the quaternion algebra $\mathbb{H}(\mathbb{Q}(\sqrt{5}))$ called the icosian ring. Its powerful arithmetic structure is presented and used to analyse its relation to the root lattice A_4. We proceed with the introduction of primitivity for icosians as well as for sublattices of A_4. Finally, we provide some tools concerning the relation of the factorisation in the icosian ring and $\mathbb{Z}[\tau]$.

Chapter 2 analyses the SSLs of the root lattice A_4 while Chapter 3 deals with its CSLs. In both chapters we derive a Dirichlet series generating function which gives the number of SSLs, respectively CSLs, of each index. In Chapter 2 we establish a parametrisation of primitive SSLs by primitive icosians. This provides the basis of the derivation of the Dirichlet series for the SSLs as well as the CSLs in Chapter 3.

In the first chapter of Part II, Chapter 4, we introduce the basic notation and definitions for powerfree words and morphisms. We proceed with a summary of relevant results for the characterisation of integer powerfree morphisms. In particular, we are interested in the question how to test a specified morphism for powerfreeness. We conclude with a section about the combinatorial entropy of powerfree words. After giving the definition, we introduce the explicit sets whose entropy is analysed in the course of this thesis. We continue with a review on how powerfree morphisms lead to lower bounds for the entropy. Two methods to derive upper bounds are introduced. The first is based on the enumeration of powerfree words of length n while the second, more efficient method relies on the central definition of *open* words and a matrix Δ_m, based on all open words of length m, whose Perron eigenvalue provides an upper bound for the entropy.

What it means for a word to be open is better understood in the context of symbolic dynamics. This leads to Chapter 5 in which powerfree words and their entropy are considered from the point of view of symbolic dynamics. Moreover, we introduce the topological entropy of a continuous map on a compact topological Hausdorff space and show that for powerfree words, as for any shift space, the combinatorial and topological entropy coincide.

In Chapter 6 we give a detailed explanation of Kolpakov's method to derive lower bounds for the entropy of powerfree words, see [**49**] for a sketch of the application of his method to three examples. We generalise his method, which starts with the Perron-Frobenius eigenvalue of the matrix Δ_m and leads, via several inductive steps, to an estimate of the number of certain power containing words. This estimation results in a procedure to calculate lower bounds for the entropy of powerfree words.

It turns out that for the examples we have analysed, very good upper and lower bounds for the entropy are achieved with the methods based on the matrix Δ_m. This matrix seems to be the central object in the investigation of the entropy of powerfree words.

We conclude with Chapter 7 where we review and apply the introduced methods to the two classical cases of ternary squarefree and binary cubefree words. For these cases we confirm Kolpakov's results from [**49**] exactly. However, for ternary minimally repetitive words our results are slightly different although our lower bound is the same. Moreover, we apply the best methods to three new cases and obtain very good upper and lower bounds for the entropy in two of the three cases. In the third we only get an upper bound while the lower bound requires a computational effort that is beyond our current computational capacities.

Part I

The Root Lattice $\mathbf{A_4}$

CHAPTER 1

Basics and Preparations

This chapter introduces the general notation and terminology of the first part of this thesis. We present a particular realisation of the A_4 root lattice in four dimensions which is motivated by the observation that it is contained in the so-called icosian ring. The powerful arithmetic structure of the icosian ring within the quaternion algebra over the real algebraic number field $\mathbb{Q}(\sqrt{5})$ is presented and the tools required in Chapter 2 and 3 are provided. We conclude with a detailed analysis of the relation of the icosian ring and the root lattice A_4.

1.1. Generalities

1.1.1. Notation. The symbols $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ denote the integer, rational, real and complex numbers, respectively. Natural numbers are always considered to be positive, i.e. $\mathbb{N} = \{1, 2, 3, \ldots\}$. If we include 0 we write $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Moreover, we set $\mathbb{R}_+ := \{\alpha \in \mathbb{R} \mid \alpha > 0\}$. The d-dimensional Euclidean space is referred to as $\mathbb{R}^d$. The Euclidean inner product of two vectors $x, y \in \mathbb{R}^d$ is denoted by

$$\langle x|y\rangle = x^t y = \sum_{i=1}^{d} x_i y_i.$$

The symbol $\subset$ is understood to include equality of sets. The cardinality of a set S is denoted by $|S|$.

For algebraic objects, e.g. groups, modules, rings, ideals, we follow the definitions in [**53**]. Let $S \subset \mathbb{R}^d$ and R be a subring of $\mathbb{R}$, then $\langle S\rangle_R$ stands for the R-hull of S. As usual, $R^\times$ denotes the group of units of a given ring R with unit. For every Abelian group G we define $G^\bullet := G \setminus \{0\}$. The direct sum of two Abelian groups G_1 and G_2 is denoted by $G_1 \oplus G_2$.

The general linear, the orthogonal and the special orthogonal group of $\mathbb{R}^d$ are referred to as $\mathrm{GL}(d)$, $\mathrm{O}(d)$ and $\mathrm{SO}(d)$ respectively. Occasionally, the corresponding groups are considered for subrings of $\mathbb{R}$, which will be clearly specified.

1.1.2. Lattices. A free $\mathbb{Z}$-module $\Gamma \subset \mathbb{R}^n$ of rank d whose $\mathbb{R}$-span is isomorphic to $\mathbb{R}^d$ is called a *d-dimensional Euclidean lattice* or *lattice* for short. Clearly, $\Gamma \subset \mathbb{R}^n$ is a lattice if and only if, there are d $\mathbb{R}$-linearly independent vectors $b_1, \ldots, b_d \in \mathbb{R}^n$, such that

$$\Gamma = \langle b_1, \ldots, b_d \rangle_{\mathbb{Z}} = \Big\{ \textstyle\sum_{i=1}^{d} m_i b_i \mid m_i \in \mathbb{Z} \Big\}. \tag{1.1}$$

The set $\{b_1, \ldots, b_d\}$ is called a *basis* of Γ. It is uniquely determined up to a matrix $Z \in \mathrm{GL}(d, \mathbb{Z})$. A matrix B_Γ whose column vectors form a basis of Γ is called a *basis matrix* of Γ and $G_\Gamma := B_\Gamma^t B_\Gamma$ is referred to as a *Gram matrix* of Γ.

Most of the time, we consider d-dimensional lattices in $\mathbb{R}^d$. So if it is not clearly specified otherwise, a lattice Γ stands for a d-dimensional lattice in $\mathbb{R}^d$. The *dual* lattice of $\Gamma \subset \mathbb{R}^d$ is defined as

$$\Gamma^* \;:=\; \{x \in \mathbb{R}^d \mid \langle x|y \rangle \in \mathbb{Z} \text{ for all } y \in \Gamma\}\,. \tag{1.2}$$

Note that $B_{\Gamma^*} = (B_\Gamma^{-1})^t$.

A lattice Γ is called *rational* if $\langle u|v \rangle \in \mathbb{Q}$ for $u, v \in \Gamma$. If $\Lambda \subset \Gamma$ is a subgroup of finite subgroup index in Γ it is called a *sublattice* of Γ. The index $[\Gamma : \Lambda]$ is defined as the number of cosets of Λ in Γ. Let us recall a helpful result from [**21**], which reveals the geometric meaning of the index, as the quotient of the volumes of the fundamental domains.

Lemma 1.1. *Let $\Gamma \subset \mathbb{R}^d$ be a lattice with basis matrix B_Γ. Λ is a sublattice of Γ if and only if there exists an invertible integer matrix Z such that $B_\Gamma Z$ is a basis matrix for Λ. The corresponding index is $[\Gamma : \Lambda] = |\det(Z)|$.*

1.1.3. Modules. For some parts of our analysis the concept of a lattice has to be generalised to free S-modules $\Gamma \subset \mathbb{R}^n$ of rank d, where S is the ring of integers of a real algebraic number field K of degree r. In other words, such a module is given by

$$\Gamma = \langle b_1, \ldots, b_d \rangle_S := \left\{ \sum_{i=1}^{d} m_i b_i \mid m_i \in S \right\} \tag{1.3}$$

where $b_1, \ldots, b_d \in \mathbb{R}^n$ are linearly independent over $\mathbb{R}$. Note that S is also a free $\mathbb{Z}$-module of rank r and thus Γ can be seen as a free $\mathbb{Z}$-module of rank rd. Basis and Gram matrices are defined analogously to the lattice case. A lattice in $\mathbb{R}^n$ can be interpreted as such an S-module with $K = \mathbb{Q}$ and thus $S = \mathbb{Z}$.

Every $\mathbb{Z}$-module is an Abelian group. Frequently, we will apply the following well known result about Abelian groups.

LEMMA 1.2. *Let Γ_1 be an Abelian group and let Γ_2 be a subgroup of Γ_1 with $[\Gamma_1 : \Gamma_2] = n$. Then, $n\Gamma_1$ is a subgroup of Γ_2.*

PROOF. Obviously, $n\Gamma_1$ is an Abelian group. For every $g \in \Gamma_1$, $g + \Gamma_2$ is an element of the finite factor group Γ_1/Γ_2 and generates a finite cyclic subgroup. Its order divides n by Lagrange's Theorem, see for example [**53**, Ch. 1, Proposition 2.2]. Consequently, $n(g + \Gamma_2) = \Gamma_2$, which means that $ng \in \Gamma_2$. □

1.2. Setting in Four Dimensions

1.2.1. The Root Lattice $\mathbf{A_4}$. The root lattice A_4 is usually defined as

$$\begin{aligned} A_4 :&= \{(x_1, \ldots, x_5) \in \mathbb{Z}^5 \mid x_1 + \ldots + x_5 = 0\} \\ &= \langle e_1 - e_2, e_2 - e_3, e_3 - e_4, e_4 - e_5 \rangle_{\mathbb{Z}}, \end{aligned} \tag{1.4}$$

where e_i denote the standard Euclidean basis vectors in $\mathbb{R}^5$. Clearly, the lattice A_4 lies in a 4-dimensional hyper-plane of $\mathbb{R}^5$, see for example [**24**]. Its Dynkin diagram is given in Figure 1.1. Following [**22**] we prefer a description

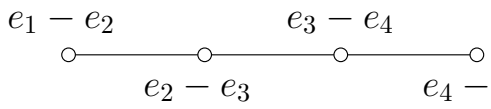


FIGURE 1.1. Standard basis representation of the root lattice A_4.

of the root lattice A_4 in $\mathbb{R}^4$, since this enables us to use the arithmetic of the quaternion algebra $\mathbb{H}(\mathbb{Q}(\sqrt{5}))$; see [**48**] for a detailed introduction to Hamilton's quaternions. Let $\tau = (1+\sqrt{5})/2$ be the golden ratio, then the lattice

$$L := \langle (1,0,0,0), \tfrac{1}{2}(-1,1,1,1), (0,-1,0,0), \tfrac{1}{2}(0,1,\tau-1,-\tau)\rangle_{\mathbb{Z}} \tag{1.5}$$

is the root lattice A_4 relative to the inner product $2\langle x|y\rangle$. The Gram matrix of L reads

$$G_L = \frac{1}{2}\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}, \tag{1.6}$$

which shows that L is a scaled copy of the root lattice A_4 in its standard representation with the basis from Figure 1.1. Note that with the lattice bases from (1.5) and Figure 1.1 the Gram matrices have the relation

$$G_{A_4} = \tfrac{1}{2} G_L. \tag{1.7}$$

1.2.2. Quaternions and the Quadratic Number field K. For brevity from now on we use the notation

$$K := \mathbb{Q}(\sqrt{5}) = \{a + b\sqrt{5} \mid a, b \in \mathbb{Q}\}, \tag{1.8}$$

which is a real quadratic number field, see [**38**] for a detailed analysis. The quaternion algebra $\mathbb{H}(K)$ is explicitly given as $\mathbb{H}(K) = K \oplus \mathrm{i}K \oplus \mathrm{j}K \oplus \mathrm{k}K$,

where the generating elements satisfy Hamilton's relations

$$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1.$$

Since $\mathbb{H}(K)$ and K^4 are isomorphic as K-vector spaces we identify the quaternion $a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d$ with the row vector (a, b, c, d). $\mathbb{H}(K)$ is equipped with a *conjugation* $\bar{\cdot}$ which is the unique map that fixes the elements of the centre of the algebra K and reverses the sign on its complement. If we write $q = (a, b, c, d) = a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d$, this means

$$\overline{q} = (a, -b, -c, -d). \tag{1.9}$$

Note that for $p, q \in \mathbb{H}(K)$, we have $\overline{pq} = \overline{q}\,\overline{p}$.

The reduced norm and trace of $q = (q_1, q_2, q_3, q_4) \in \mathbb{H}(K)$ are defined by

$$\mathrm{nr}(q) := q\overline{q} = \sum_{i=1}^{4} q_i^2 = \overline{q}q \quad \text{and} \quad \mathrm{tr}(q) := q + \overline{q} = 2q_1 \tag{1.10}$$

where we canonically identify an element $\alpha \in K$ with the quaternion $(\alpha, 0, 0, 0)$. For any $q \in \mathbb{H}(K)$, $|q|$ denotes its Euclidean length, which need not be an element of K. Nevertheless, one has $|rs| = |r||s|$ for arbitrary $r, s \in \mathbb{H}(K)$. Due to the geometric meaning, we use the notations $|q|^2$ and $\mathrm{nr}(q)$ in parallel. Obviously, the inverse of $q \in \mathbb{H}(K)^\bullet$ is given by

$$q^{-1} = \frac{1}{\mathrm{nr}(q)}\overline{q}.$$

Note that for $p, q \in \mathbb{H}(K)$ the identity

$$\mathrm{nr}(p + q) = \mathrm{nr}(p) + \mathrm{nr}(q) + \mathrm{tr}(p\overline{q}), \tag{1.11}$$

holds, which shows that nr is a quadratic form on the vector space $\mathbb{H}(K)$. Moreover, we have $\mathrm{nr}(pq) = \mathrm{nr}(p)\,\mathrm{nr}(q)$, i.e. nr is multiplicative, and $\mathrm{tr}(pq) = \mathrm{tr}(qp)$.

An element $q \in \mathbb{H}(K)$ is called *integral* when both $\mathrm{nr}(q)$ and $\mathrm{tr}(q)$ are elements of

$$\mathbb{Z}[\tau] := \{m + n\tau \mid m, n \in \mathbb{Z}\}, \tag{1.12}$$

which is the ring of integers of the quadratic field K, see [**38**] for a detailed analysis. Here, we only recall the properties which are needed later.

The algebraic conjugation in K, as determined by the map $\sqrt{5} \mapsto -\sqrt{5}$, is denoted by $'$. For $\alpha = a + b\tau \in K$ we define its absolute norm as well as its trace as

$$\mathrm{N}(\alpha) := |\alpha\alpha'| = |(a + b\tau)(a + b\tau')| = |a^2 + ab - b^2| \tag{1.13}$$

$$\mathrm{Tr}(\alpha) := \alpha + \alpha' = 2a + b. \tag{1.14}$$

Clearly, $\alpha \in \mathbb{Z}[\tau]$ implies $\mathrm{N}(\alpha) \in \mathbb{N}$. Moreover, the absolute norm is multiplicative, i.e. for $\alpha, \beta \in K$ we have $\mathrm{N}(\alpha\beta) = \mathrm{N}(\alpha)\,\mathrm{N}(\beta)$. The units of $\mathbb{Z}[\tau]$ are

$$\mathbb{Z}[\tau]^\times := \{\alpha \in \mathbb{Z}[\tau] \mid \mathrm{N}(\alpha) = 1\} = \{\pm\tau^n \mid n \in \mathbb{Z}\}\,; \tag{1.15}$$

see [**38**, Theorem 257]. An element $\alpha \in \mathbb{Z}[\tau]$ is called an *associate* of $\beta \in \mathbb{Z}[\tau]$, if $\alpha\beta^{-1} \in \mathbb{Z}[\tau]^\times$. Since K is a Euclidean field, see [**38**, Theorem 247], the ring $\mathbb{Z}[\tau]$ is a principal ideal domain. Hence every ideal $\mathfrak{a}$ in $\mathbb{Z}[\tau]$ has the form $\mathfrak{a} = \alpha\mathbb{Z}[\tau]$ for some $\alpha \in \mathbb{Z}[\tau]$. If $\alpha\mathbb{Z}[\tau]$ is a non-zero ideal, its index, i.e. the number of cosets of $\alpha\mathbb{Z}[\tau]$ in $\mathbb{Z}[\tau]$, is given by

$$[\mathbb{Z}[\tau] : \alpha\mathbb{Z}[\tau]] = \mathrm{N}(\alpha) = |\alpha\alpha'|,$$

compare [**34**, Theorem 2.7]. The Dirichlet series generating function for the number of non-zero ideals of a given index is the Dedekind zeta function of the algebraic number field K, see [**87**]. It reads

$$\zeta_K(s) = \sum_{\mathfrak{a} \subset \mathbb{Z}[\tau]} \frac{1}{[\mathbb{Z}[\tau] : \mathfrak{a}]^s} = \sum_{m=1}^{\infty} \frac{a_K(m)}{m^s},$$

where $\mathfrak{a}$ runs through the non-zero ideals of $\mathbb{Z}[\tau]$ and $a_K(m)$ denotes the number of ideals of index m. The prime numbers of $\mathbb{Z}[\tau]$ are characterised as follows, see [**38**, Theorem 257].

THEOREM 1.3. *Let $\pi = (m + n\tau) \in \mathbb{Z}[\tau]$, such that $\mathrm{N}(\pi) = p$ is a prime number of $\mathbb{Z}$. Then, the associates of the following elements are the prime numbers of $\mathbb{Z}[\tau]$:*

(1) $\sqrt{5}$
(2) *p, if $p \equiv \pm 2 \mod 5$*
(3) *$(m + n\tau)$ and $(m + n\tau')$ with $(m + n\tau)(m + n\tau') = p$, if $p \equiv \pm 1 \mod 5$.*

□

The prime numbers of the first, second and third type, are called *ramified, inert* and *splitting* primes, respectively. They lead to the following explicit expression of the zeta function

(1.16)
$$\zeta_K(s) = \frac{1}{1-5^{-s}} \prod_{p \equiv \pm 1(5)} \frac{1}{(1-p^{-s})^2} \prod_{p \equiv \pm 2(5)} \frac{1}{1-p^{-2s}}$$
$$= 1 + \tfrac{1}{4^s} + \tfrac{1}{5^s} + \tfrac{1}{9^s} + \tfrac{2}{11^s} + \tfrac{1}{16^s} + \tfrac{2}{19^s} + \tfrac{1}{20^s} + \tfrac{1}{25^s} + \tfrac{2}{29^s} + \dots,$$

see [**12**] for further details including asymptotic properties.

1.2.3. Orders and Ideals. We call a finitely generated $\mathbb{Z}[\tau]$-module $\mathcal{I} \subset \mathbb{H}(K)$ with $\langle \mathcal{I} \rangle_K = \mathbb{H}(K)$, a *full $\mathbb{Z}[\tau]$-module* (in $\mathbb{H}(K)$). Note that in [**85, 13**] these modules are called 'ideals'. Here, the term 'full $\mathbb{Z}[\tau]$-module' is used instead, in order to avoid confusion with ideals from ring theory.

An *order* is a full $\mathbb{Z}[\tau]$-module that is also a subring of $\mathbb{H}(K)$ containing 1. It is called *maximal* when it is not contained in any larger order which is not $\mathbb{H}(K)$, see for example [**85**, Chap. 1, Sec. 4] and [**74**] for details.

Clearly, if $q \in \mathbb{H}(K)$ is integral then so is $\overline{q}$. However, the set of all integral quaternions fails to be a ring due to the fact that quaternion multiplication is non-commutative, see for example [**85, 79**]. Hence instead

of considering all integral quaternions, orders of integral quaternions are considered. From [**13**, Fact 3] we recall

LEMMA 1.4. *All elements of an arbitrary order $\mathcal{O} \subset \mathbb{H}(K)$ are integral. Moreover, one has $\overline{\mathcal{O}} = \mathcal{O}$ and $\mathcal{O} \cap K = \mathbb{Z}[\tau]$.* □

Obviously,

$$\mathcal{L} := \langle 1, \mathrm{i}, \mathrm{j}, \mathrm{k} \rangle_{\mathbb{Z}[\tau]} \;\cong\; \mathbb{Z}[\tau]^4 \tag{1.17}$$

is an order, as is $q\mathcal{L}q^{-1}$ for every $q \in \mathbb{H}(K)^\bullet$, but it is not maximal since appending the integral quaternion $\frac{1}{2}(1,1,1,1)$ generates a larger order.

Let $\mathcal{O}$ be an order. A *left $\mathbb{Z}[\tau]$-module of $\mathcal{O}$* is a full $\mathbb{Z}[\tau]$-module $\mathcal{I}$ with $q\mathcal{I} \subset \mathcal{I}$ for every $q \in \mathcal{O}$. It is called *principal* if there is a $q \in \mathbb{H}(K)$ such that $\mathcal{I} = \mathcal{O}q$. Note that every left ideal of $\mathcal{O}$ is a left $\mathbb{Z}[\tau]$-module of $\mathcal{O}$. Right $\mathbb{Z}[\tau]$-modules of $\mathcal{O}$ and principal right $\mathbb{Z}[\tau]$-module of $\mathcal{O}$ are defined accordingly. From now on we speak of *one-sided $\mathbb{Z}[\tau]$-modules*, if we mean either left or right $\mathbb{Z}[\tau]$-modules.

The *class number* of $\mathcal{O}$ is the number of equivalence classes of left $\mathbb{Z}[\tau]$-modules of $\mathcal{O}$ under the following equivalence relation: Let $\mathcal{I}$ and $\mathcal{J}$ be two left $\mathbb{Z}[\tau]$-modules of $\mathcal{O}$. They are equivalent, if there is an $a \in \mathbb{H}(K)^\bullet$, such that $\mathcal{J} = \mathcal{I}a$.

Since the conjugation of $\mathbb{H}(K)$ maps right multiplication to left multiplication and $\mathcal{O}$ to itself, interchanging the roles of 'left' and 'right' results in the same value of the class number.

From [**13**, Fact 5] we know that all maximal orders of $\mathbb{H}(K)$ have the same class number, which is known as the class number of $\mathbb{H}(K)$. Moreover, if the class number is 1, all maximal orders are mutual images of one another under inner automorphisms, i.e. under maps of the form $f : \mathbb{H}(K) \to \mathbb{H}(K), x \mapsto axa^{-1}$ for $a \in \mathbb{H}(K)^\bullet$. In [**85**, p. 156] and in [**79**, Ap. A.3] we find that the class number of $\mathbb{H}(K)$ is 1. Hence, all one-sided $\mathbb{Z}[\tau]$-modules of a maximal order $\mathcal{O}$ are equivalent.

This means that for every left ideal $\mathcal{I}$ of a maximal order $\mathcal{O}$ there is a $q \in \mathcal{O}^\bullet$ such that $\mathcal{I} = \mathcal{O}q$, since $\mathcal{O}$ itself is obviously a left ideal of $\mathcal{O}$. Of course these arguments work for right ideals analogously. Hence every one-sided ideal of $\mathcal{O}$ is principal.

1.2.4. $\mathbb{Z}[\tau]$-Indices. Following [**13**] we denote the determinant of an endomorphisms φ of the K-vector space $\mathbb{H}(K)$ by $\det_K(\varphi)$. It is an element of K and can be calculated using any K-basis of $\mathbb{H}(K)$, as it is not dependent on the particular choice of the basis.

If $\mathcal{J} \subset \mathcal{I}$ are full $\mathbb{Z}[\tau]$-modules in $\mathbb{H}(K)$, we define the K-index of $\mathcal{J}$ in $\mathcal{I}$ as

$$[\mathcal{I} : \mathcal{J}]_K := \det_K(\varphi), \tag{1.18}$$

where φ is a linear map that takes a $\mathbb{Z}[\tau]$-basis of $\mathcal{I}$ to a $\mathbb{Z}[\tau]$-basis of $\mathcal{J}$. The K-index is well-defined up to units of $\mathbb{Z}[\tau]$, as changing the bases of $\mathcal{I}$ and $\mathcal{J}$ multiplies it by an element of $\mathbb{Z}[\tau]^\times$. Since $\mathcal{J} \subset \mathcal{I}$, we know that $\det_K(\varphi) \in \mathbb{Z}[\tau]$. By considering the inverse of φ it is clear that $\mathcal{I} = \mathcal{J}$ if and only if $[\mathcal{I} : \mathcal{J}]_K \in \mathbb{Z}[\tau]^\times$. If $\mathcal{I}_1 \subset \mathcal{I}_2 \subset \mathcal{I}_3$ are full $\mathbb{Z}[\tau]$-modules in $\mathbb{H}(K)$, then

$$[\mathcal{I}_3 : \mathcal{I}_1]_K = [\mathcal{I}_3 : \mathcal{I}_2]_K [\mathcal{I}_2 : \mathcal{I}_1]_K,$$

i.e. the K-index is multiplicative. The following lemma is analogous to the lattice case, see Lemma 1.2, and its claim is obvious with the definition of the K-index in (1.18).

LEMMA 1.5. *Let $\mathcal{I}$ be a full $\mathbb{Z}[\tau]$-module in $\mathbb{H}(K)$ with basis matrix $B_\mathcal{I}$. Then, $\mathcal{J}$ is a full $\mathbb{Z}[\tau]$-submodule of $\mathcal{I}$ if and only if there exists a non-singular matrix Z with $\mathbb{Z}[\tau]$-entries such that $B_\mathcal{I} Z = B_\mathcal{J}$ is a basis matrix for $\mathcal{J}$. The corresponding index is then $[\mathcal{I} : \mathcal{J}]_K = \det(Z)$.* □

In complete analogy to the lattice case, compare [**3**, Lemma 2.2], this description of full $\mathbb{Z}[\tau]$-submodules leads to

LEMMA 1.6. *Let $\mathcal{J} \subset \mathcal{I}$ be full $\mathbb{Z}[\tau]$-modules in $\mathbb{H}(K)$ with $[\mathcal{I} : \mathcal{J}]_K = \alpha$. Then, $\alpha\mathcal{I} \subset \mathcal{J}$ is a full $\mathbb{Z}[\tau]$-module of index α^3.*

PROOF. By Lemma 1.5 there is a matrix Z with $\mathbb{Z}[\tau]$-entries such that $B_{\mathcal{I}}Z = B_{\mathcal{J}}$ and $[\mathcal{I} : \mathcal{J}]_K = \det(Z) = \alpha$. The standard formula for the inverse matrix implies that αZ^{-1} is a non-singular matrix with $\mathbb{Z}[\tau]$-entries. So again by Lemma 1.5, $\alpha B_{\mathcal{J}} Z^{-1} = \alpha B_{\mathcal{I}}$ which implies that $\alpha\mathcal{I}$ is a full $\mathbb{Z}[\tau]$-submodule of $\mathcal{J}$ whose index is $\det(\alpha Z^{-1}) = \alpha^4 \frac{1}{\det(Z)} = \alpha^3$. □

From [**13**, Fact 8] we recall the following lemma, which describes the relationship between the K-index and the usual coset-counting index.

LEMMA 1.7. *Let $\mathcal{I}, \mathcal{J}$ be two full $\mathbb{Z}[\tau]$-modules in $\mathbb{H}(K)$ with $\mathcal{J} \subset \mathcal{I}$. Then,*

$$[\mathcal{I} : \mathcal{J}] = \mathrm{N}([\mathcal{I} : \mathcal{J}]_K).$$

□

We conclude this subsection with two lemmas that simplify the calculation of certain indices. The first can either be found in [**13**, Section 5, Lemma 1] or its claim is deduced by means of straightforward calculations with the basis matrices according to Lemma 1.5.

LEMMA 1.8. *For any order $\mathcal{O}$ in $\mathbb{H}(K)$ and any $q \in \mathcal{O}^\bullet$, one has*

$$[\mathcal{O} : q\mathcal{O}]_K = [\mathcal{O} : \mathcal{O}q]_K = \mathrm{nr}(q)^2 = |q|^4.$$

LEMMA 1.9. *Let $\mathcal{O}$ be an order in $\mathbb{H}(K)$ and let $\mathcal{J} \subset \mathcal{O}$ be a full $\mathbb{Z}[\tau]$-module of $\mathbb{H}(K)$. For $a, b \in \mathcal{O}^\bullet$ one has*

$$[\mathcal{O} : a\mathcal{J}b] = \mathrm{N}(\mathrm{nr}(a)^2 \,\mathrm{nr}(b)^2)[\mathcal{O} : \mathcal{J}].$$

PROOF. Obviously, we have $a\mathcal{J}b \subset a\mathcal{O}b \subset a\mathcal{O} \subset \mathcal{O}$ and

$$[a\mathcal{O}b : a\mathcal{J}b] = [\mathcal{O} : J].$$

Moreover, a combination of Lemma 1.8 and 1.7 gives $[a\mathcal{O} : a\mathcal{O}b] = [\mathcal{O} : \mathcal{O}b] = \mathrm{N}(\mathrm{nr}(b)^2)$ and $[\mathcal{O} : a\mathcal{O}] = \mathrm{N}(\mathrm{nr}(a)^2)$. Hence due to the multiplicativity of the index we have

$$[\mathcal{O} : a\mathcal{J}b] = [\mathcal{O} : a\mathcal{O}][a\mathcal{O} : a\mathcal{O}b][a\mathcal{O}b : a\mathcal{J}b] = \mathrm{N}(\mathrm{nr}(a)^2)\,\mathrm{N}(\mathrm{nr}(b)^2)[\mathcal{O} : \mathcal{J}].$$

□

1.2.5. The Icosian Ring. In this subsection we introduce a particular maximal order of $\mathbb{H}(K)$, the icosian ring, and summarise its properties as far as they are required in Chapter 2 and 3; for more details including its connection to the root system of type H_4 see for example [**60, 59, 22, 12, 24**] and references given there.

Following [**24**, Ch. 8, Sec. 2.1] we define the *icosian group* I as the multiplicative group of order 120 consisting of the quaternions

$$(\pm 1, 0, 0, 0)\,,\ \tfrac{1}{2}(\pm 1, \pm 1, \pm 1, \pm 1)\,,\ \tfrac{1}{2}(\pm\tau', \pm\tau, 0, \pm 1) \tag{1.19}$$

and all their *even* coordinate permutations. The *icosian ring* $\mathbb{I}$ is defined as $\mathbb{I} := \langle I\rangle_{\mathbb{Z}}$, and has rank 8 as a $\mathbb{Z}$-module. At the same time, $\mathbb{I}$ is a $\mathbb{Z}[\tau]$-module of rank 4, and can alternatively be written as

$$\mathbb{I} := \Big\langle (1,0,0,0), (0,1,0,0), \tfrac{1}{2}(1,1,1,1), \tfrac{1}{2}(1-\tau, \tau, 0, 1)\Big\rangle_{\mathbb{Z}[\tau]}. \tag{1.20}$$

The elements of $\mathbb{I}$ are called *icosians*.

From [**60**] we recall that the icosian group I is characterised as follows within $\mathbb{I}$:

$$I = \Big\{ q \in \mathbb{I} \mid |q|^2 = 1 \Big\}. \tag{1.21}$$

Recall that $'$ stands for the algebraic conjugation in K. A straightforward calculation with the basis matrix resulting from (1.20) reveals that $\mathbb{I} \neq \mathbb{I}'$ and that $\mathbb{I}$ as well as $\mathbb{I}'$ are maximal orders of $\mathbb{H}(K)$, in fact they are the only maximal orders that contain $\mathcal{L}$ of (1.17); compare for example [**13**]. By

means of Lemmas 1.5 and 1.7 we see that

$$[\mathbb{I} : \mathcal{L}] = 16 = [\mathbb{I}' : \mathcal{L}].$$

Moreover, with (1.20) it is clear that $2\mathbb{I} \subset \mathcal{L}$.

Note that $(1+i)\mathbb{I} = \mathbb{I}'(1+i)$, so that $(1+i)\mathbb{I}$ is a one-sided, but not a two-sided ideal in $\mathbb{I}$. This relation also shows that $\mathbb{I}$ and $\mathbb{I}'$ are related by an inner automorphism of $\mathbb{H}(K)$.

Considering Lemma 1.4 with respect to $\mathbb{I}$ as maximal order, it is obvious that $q \in \mathbb{I}$ with $\mathrm{nr}(q) \in \mathbb{Z}[\tau]^\times$ implies that $q^{-1} = \frac{\bar{q}}{\mathrm{nr}(q)} \in \mathbb{I}$. Conversely, if $q \in \mathbb{I}^\times$, we know that $\mathrm{nr}(q)$ and $\mathrm{nr}(q^{-1}) = \mathrm{nr}(q)^{-1} \in \mathbb{Z}[\tau]$, i.e. $\mathrm{nr}(q) \in \mathbb{Z}[\tau]^\times$. With (1.15) this leads to

(1.22)
$$\mathbb{I}^\times \;=\; \{q \in \mathbb{I} \mid \mathrm{N}(|q|^2) = 1\} = \{\pm\tau^n q \mid n \in \mathbb{Z} \text{ and } q \in I\} \;\cong\; \mathbb{Z}[\tau]^\times \times I.$$

The Dirichlet series generating function for the number of non-zero one-sided-ideals contained in $\mathbb{I}$ of a given index reads

$$\begin{aligned}\zeta_{\mathbb{I}}(s) &= \sum_{\mathcal{I}\subset\mathbb{I}} \frac{1}{[\mathbb{I}:\mathcal{I}]^s} = \sum_{m=1}^{\infty} \frac{a_{\mathbb{I}}(m)}{m^{2s}} = \zeta_K(2s)\zeta_K(2s-1) \\ &= 1 + \tfrac{5}{16^s} + \tfrac{6}{25^s} + \tfrac{10}{81^s} + \tfrac{24}{121^s} + \tfrac{21}{256^s} + \tfrac{40}{361^s} + \tfrac{30}{400^s} + \tfrac{31}{625^s} + \dots\end{aligned} \tag{1.23}$$

where $\mathcal{I}$ runs through the non-zero one-sided ideals of $\mathbb{I}$, see [**12**] for its derivation and further details including asymptotic properties. The arithmetic function $a_{\mathbb{I}}(m)$ is multiplicative, and hence completely specified by its values at prime powers. These are given by

$$a_{\mathbb{I}}(p^r) \;=\; \begin{cases} \frac{5^{r+1}-1}{4}, & \text{if } p = 5, \\ \sum_{\ell=0}^{r} (\ell+1)(r-\ell+1)p^\ell, & \text{for primes } p \equiv \pm 1 \ (5), \\ \frac{1-p^{r+2}}{1-p^2}, & \text{for primes } p \equiv \pm 2 \ (5) \text{ and } r \text{ even}, \\ 0, & \text{for primes } p \equiv \pm 2 \ (5) \text{ and } r \text{ odd}. \end{cases} \tag{1.24}$$

Note that by Lemma 1.9, we know that

$$[\mathbb{I} : q\mathbb{I}] = \mathrm{N}(\mathrm{nr}(q))^2 = [\mathbb{I} : \mathbb{I}q]. \tag{1.25}$$

Hence the possible indices, here the denominators, are the squares of integers that are representable by the quadratic form $x + xy - y^2$.

The Dirichlet series generating function, which counts the number of non-zero two-sided-ideals of $\mathbb{I}$, is of an even easier form. It reads $\zeta_{\mathbb{I}.\mathbb{I}}(s) = \zeta_K(4s)$, see [**13**, Eq. (19)] . Now, we easily conclude that the two-sided ideals in $\mathbb{I}$ are of a very convenient type.

LEMMA 1.10. *The two-sided ideals in $\mathbb{I}$ have the form $\alpha\mathbb{I}$ with $\alpha \in \mathbb{Z}[\tau]^\bullet$.*

PROOF. Of course, for every $\alpha \in \mathbb{Z}[\tau]$ the ideal $\alpha\mathbb{I} = \mathbb{I}\alpha$ is two-sided, and $[\mathbb{I} : \alpha\mathbb{I}] = \mathrm{N}(\alpha)^4$ by Lemma 1.9. Since $\zeta_{\mathbb{I}.\mathbb{I}}(s) = \zeta_K(4s)$ this means that these are already all two-sided ideals in $\mathbb{I}$. □

LEMMA 1.11. *For every $\alpha \in \mathbb{Z}[\tau]$, there is an icosian q with $\alpha = \mathrm{nr}(q) = q\overline{q}$. Moreover, if $\alpha \in \mathbb{Z}[\tau]^\times$, then q and $\overline{q}$ are associated.*

PROOF. Since the reduced norm nr is multiplicative it is sufficient to show the following for the first claim: For every prime number $\pi \in \mathbb{Z}[\tau]$, there is an icosian $q \in \mathbb{I}$ with $\pi = \mathrm{nr}(q) = q\overline{q}$.

If we assume that there is no icosian q with $\mathrm{nr}(q) = \pm\sqrt{5}$, this implies, due to $[\mathbb{I} : \mathbb{I}q] = \mathrm{N}(\pm\sqrt{5})^2$, compare (1.25), that $a_{\mathbb{I}}(5) = 0$, which contradicts (1.24). Thus there is an icosian q with $\mathrm{nr}(q) = \pm\sqrt{5}$. Similarly, for $p \in \mathbb{Z}$ with $p \equiv \pm 2 \mod 5$ there is an icosian q with $\mathrm{nr}(q) = p$, since $[\mathbb{I} : \mathbb{I}q] = p^4$ and $a_{\mathbb{I}}(p^2) > 0$. For the remaining case of the splitting prime $p \in \mathbb{Z}$ with $p \equiv \pm 1 \mod 5$ and $p = \pi\pi'$ note that there exists an icosian q with either $\mathrm{nr}(q) = \pi$ or $\mathrm{nr}(q) = \pi'$, as $a_{\mathbb{I}}(p) > 0$. If $\mathrm{nr}(q) = \pi$ then $\mathrm{nr}(\widetilde{q}) = \pi'$ or vice versa. This gives the first claim.

For the second claim note that if $\mathrm{nr}(q) = \alpha \in \mathbb{Z}[\tau]^\times$, we know that $q(\overline{q})^{-1} = \frac{1}{\mathrm{nr}(q)}q^2 \in \mathbb{I}$ and considering (1.22) this means that $q(\overline{q})^{-1} \in \mathbb{I}^\times$. This shows that q and $\overline{q}$ are associated. □

If we substitute in the proof of Lemma 1.11 the arithmetic functions $a_{\mathbb{I}}$ by the arithmetic function which counts the number of ideals in $\mathbb{I}$, generated by a primitive icosian, see (2.8) below, we get

COROLLARY 1.12. *For every $\alpha \in \mathbb{Z}[\tau]$, there is a primitive icosian p with $\alpha = \mathrm{nr}(p) = p\overline{p}$. Moreover, p and $\overline{p}$ are not associated if and only if $\alpha \notin \mathbb{Z}[\tau]^{\times}$.*

PROOF. It only remains to show the second statement. Note that p and $\overline{p}$ are associated if and only if $p(\overline{p})^{-1} = \frac{1}{\mathrm{nr}(p)}p^2 \in \mathbb{I}$, which is equivalent to $\mathrm{nr}(p) \in \mathbb{Z}[\tau]^{\times}$, due to the primitivity of p^2. □

One can use the quadratic form defined by $\mathrm{tr}(x\overline{y}) = 2\langle x|y\rangle$ to define the *dual* of a full $\mathbb{Z}[\tau]$-module $\mathcal{I} \subset \mathbb{H}(K)$ as

$$\mathcal{I}^* = \{x \in \mathbb{H}(K) \mid \mathrm{tr}(x\overline{y}) \in \mathbb{Z}[\tau] \text{ for all } y \in \mathcal{I}\}. \tag{1.26}$$

With this definition, one has the following important property of the icosian ring, compare [**60**] for details.

LEMMA 1.13. *The icosian ring is self-dual, i.e. one has $\mathbb{I}^* = \mathbb{I}$.*

1.3. Characterisation and Factorisation

Our choice of the realisation of the root lattice A_4 in form of the lattice L from (1.5) is particularly motivated by the observation that $L \subset \mathbb{I}$. In this section this inclusion is analysed in more detail.

1.3.1. The Module $L[\tau]$. The given basis of the lattice L in (1.5) generates the full $\mathbb{Z}[\tau]$-module

$$L[\tau] := \left\langle (1,0,0,0), \tfrac{1}{2}(-1,1,1,1), (0,-1,0,0), \tfrac{1}{2}(0,1,\tau-1,-\tau)\right\rangle_{\mathbb{Z}[\tau]} \tag{1.27}$$

which is a submodule of $\mathbb{I}$. A straightforward calculation with Lemmas 1.5 and 1.7 reveals that

$$[\mathbb{I} : L[\tau]] = 5. \tag{1.28}$$

Since L is a rational lattice and τL obviously not, we have $L \cap \tau L = \{0\}$, so that

$$L[\tau] = L \oplus \tau L. \tag{1.29}$$

1.3.2. The Twist Map. The detailed arithmetic structure of $\mathbb{I}$ is the key to the characterisation of the similar sublattices and coincidence site lattices for L in Chapter 2 and 3. Another important object is the following map, called the *twist map*, which is an involution of the second kind for $\mathbb{H}(K)$; see [**47**] for details. If $q = (a, b, c, d)$, it is defined by the mapping

$$q \mapsto \widetilde{q} := (a', b', d', c'). \tag{1.30}$$

The relevance of this rather strange looking map, with its combination of a permutation of two coordinates with algebraic conjugation of all coordinates, was noticed in [**39**] as a result of explicit calculations. It has the following important properties.

LEMMA 1.14. *The twist map* $\widetilde{\ }$ *of* (1.30) *is a* $\mathbb{Q}$*-linear and* K*-semi-linear involutory algebra anti-automorphism, i.e. for arbitrary* $p, q \in \mathbb{H}(K)$ *and* $\alpha \in K$*, it satisfies:*

(a) $\widetilde{p+q} = \widetilde{p} + \widetilde{q}$ *and* $\widetilde{\alpha p} = \alpha' \widetilde{p}$,
(b) $\widetilde{pq} = \widetilde{q}\,\widetilde{p}$ *and* $\widetilde{\widetilde{p}} = p$,
(c) $\widetilde{\overline{p}} = \overline{\widetilde{p}}$ *and thus, for* $p \neq 0$*, also* $(\widetilde{p})^{-1} = \widetilde{p^{-1}}$.

It maps K*, the centre of the algebra* $\mathbb{H}(K)$*, onto itself, but fixes only the elements of* $\mathbb{Q}$ *within* K*. Moreover,* $\widetilde{\mathbb{I}} = \mathbb{I}$ *and* $\widetilde{\mathbb{I}'} = \mathbb{I}'$.

PROOF. Most of these properties are immediate from the definition, and (b) can be proved by checking the action of the twist map on the basis quaternions $1, \mathrm{i}, \mathrm{j}, \mathrm{k}$. The statements on K, $\mathbb{I}$ and $\mathbb{I}'$ follow easily from the definition in (1.20). □

Seen as a vector space over $\mathbb{Q}$, $\mathbb{H}(K)$ has dimension 8 and can be split into a direct sum, $\mathbb{H}(K) = V_+ \oplus V_-$, where

$$V_\pm := \{x \in \mathbb{H}(K) \mid \widetilde{x} = \pm x\}$$

are the eigenspaces under the twist map, with $V_+ \cap V_- = \{0\}$ and $V_- = \sqrt{5}\, V_+$. Note that $L \subset V_+$, so the four basis vectors of L from (1.5) form a $\mathbb{Q}$-basis of V_+. This observation combined with the nature of $\mathbb{I}$ as a maximal order in $\mathbb{H}(K)$ suggests that L might actually be the $\mathbb{Z}$-module of fixed points of the twist map inside $\mathbb{I}$.

PROPOSITION 1.15. *The lattice L can be characterised in $\mathbb{I}$ as follows:*

(i) $L = \{x \in \mathbb{I} \mid \widetilde{x} = x\} = \mathbb{I} \cap V_+$

(ii) $L = \{x + \widetilde{x} \mid x \in \mathbb{I}\} = \varphi_+(\mathbb{I})$, *where the $\mathbb{Q}$-linear map* $\varphi_+ : \mathbb{H}(K) \to \mathbb{H}(K)$, *is defined by* $\varphi_+(x) = x + \widetilde{x}$.

PROOF. Note that the basis vectors of L in (1.5) are fixed under the twist map, so $\widetilde{x} = x$ is clear for all $x \in L$. We also know from Lemma 1.14 that $\widetilde{\mathbb{I}} = \mathbb{I}$. To prove our claim, we have to show that no element of $\mathbb{I} \setminus L$ is fixed under the twist map. Recall from (1.28) that $[\mathbb{I} : L[\tau]] = 5$, so that $\mathbb{I}/L[\tau] \simeq C_5$. Consequently, the twist map, which is an involution, induces an automorphism on the cyclic group C_5. The order of this automorphism must divide 2. With the involved basis matrices $B_{\mathbb{I}}$ and $B_{L[\tau]}$ it is easily checked that,

$$0 \neq v = \mathrm{j} - \mathrm{k} \in \mathbb{I} \setminus L[\tau].$$

Since $\widetilde{v} = -v$, the induced automorphism on C_5 cannot be the identity. This leaves only inversion (i.e. $k \mapsto -k \bmod 5$), which has no fixed point in C_5 other than 0, so that all fixed points of the twist map inside $\mathbb{I}$ must lie in $L[\tau]$. It is easy to see that a quaternion of the form $x + \tau y$ with $x, y \in L$ is fixed if and only if $y = 0$, which completes the argument.

The characterisation (i) implies that $x + \widetilde{x} \in L$ for every $x \in \mathbb{I}$. On the other hand, observing that $\tau' = 1 - \tau$, any $x \in L$ permits the decomposition

$$x \;=\; (\tau + \tau')\, x \;=\; \tau x + \tau' \widetilde{x} \;=\; \tau x + \widetilde{\tau x}\,.$$

Since x is an element of the $\mathbb{Z}[\tau]$-module $\mathbb{I}$, we know that $\tau x \in \mathbb{I}$ and the claim follows. □

For our further analysis it is convenient to have several ways of characterising the submodule $L[\tau]$ within the icosian ring $\mathbb{I}$.

Proposition 1.16. *The submodule $L[\tau]$ of $\mathbb{I}$ can be characterised as follows:*

$$L[\tau] = \Big\{ x \in \mathbb{I} \mid (x - \widetilde{x}) \in \sqrt{5}\,\mathbb{I} \Big\} = \Big\{ x \in \mathbb{I} \mid 5 \text{ divides } |x - \widetilde{x}|^2 \Big\} \tag{1.31}$$

Proof. Every quaternion $x \in L[\tau]$ can be written as $x = a + \tau b$ with $a, b \in L$; see (1.29). Using $a = \widetilde{a}$ and $b = \widetilde{b}$, we see that for all $q \in L[\tau]$

$$x - \widetilde{x} = a + \tau b - a - (1 - \tau)b = (2\tau - 1)b \in (2\tau - 1)L \subset (2\tau - 1)\mathbb{I} = \sqrt{5}\,\mathbb{I}.$$

On the other hand a straightforward calculation with the involved basis matrices reveals that every $x \in \mathbb{I}$ can be written as $x = q + ku$, where $q \in L[\tau]$, $u = \frac{1}{2}(1 - \tau, \tau, 0, 1)$ and $0 \le k \le 4$. Clearly,

$$u - \widetilde{u} = \tfrac{1}{2}(1 - 2\tau, 2\tau - 1, -1, 1) \notin \sqrt{5}\,\mathbb{I}$$

and hence $x - \widetilde{x} = (q - \widetilde{q}) + k(u - \widetilde{u}) \in \sqrt{5}\,\mathbb{I}$ implies that $k = 0$. This means that $x \in L[\tau]$, which proves the first equality of (1.31). To prove the second equality observe that

$$\begin{aligned} |x - \widetilde{x}|^2 &= |q - \widetilde{q}|^2 + k(q - \widetilde{q})\overline{(u - \widetilde{u})} + k(u - \widetilde{u})\overline{(q - \widetilde{q})} + k^2 |u - \widetilde{u}|^2 \\ &= |q - \widetilde{q}|^2 + \mathrm{tr}(k(q - \widetilde{q})\overline{(u - \widetilde{u})}) + 12k^2. \end{aligned}$$

By the first equality (1.31) it is clear that $\sqrt{5}$ divides $|q-\widetilde{q}|^2$ and $\mathrm{tr}(k(q-\widetilde{q})\overline{(u-\widetilde{u})})$. Hence only if $k=0$, i.e. $x \in L[\tau]$, $|x-\widetilde{x}|^2$ is divisible by 5.

To show the remaining inclusion for the second equality of (1.31) let $x \in L[\tau]$ and let $q \in \mathbb{I}$ such that $x-\widetilde{x} = \sqrt{5}q$. This implies that $|x-\widetilde{x}|^2 = 5|q|^2$, i.e. 5 divides $|x-\widetilde{x}|^2$. □

An immediate consequence of the previous proposition is

COROLLARY 1.17. *One has* $\sqrt{5}\,\mathbb{I} = (2\tau-1)\mathbb{I} \subset L[\tau]$. □

1.3.3. Primitivity. A sublattice Λ of L is called *L-primitive* when $\alpha\Lambda \subset L$, with $\alpha \in \mathbb{Q}$, implies $\alpha \in \mathbb{Z}$. Similarly, an element $p \in \mathbb{I}$ is called *$\mathbb{I}$-primitive* when $\alpha p \in \mathbb{I}$, this time with $\alpha \in K$, is only possible with $\alpha \in \mathbb{Z}[\tau]$. For brevity, we simply use the term 'primitive' in both cases, whenever the meaning is clear from the context.

Note that an icosian p is primitive if and only if all $\mathbb{Z}[\tau]$-divisors of p are units. Due to the fact that $\mathbb{I}$ is a unique factorisation domain, this means that if p is primitive p^n is a primitive icosian, too.

Depending on the context let lcm stand for the lowest common multiple in $\mathbb{Z}$ or in $\mathbb{Z}[\tau]$. More precisely, for $\alpha, \beta \in \mathbb{Z}[\tau]$ we define, analogously to the integer case, $\mathrm{lcm}(\alpha,\beta) \in \mathbb{Z}[\tau]$ such that

$$\mathrm{lcm}(\alpha,\beta)\mathbb{Z}[\tau] = \alpha\mathbb{Z}[\tau] \cap \beta\mathbb{Z}[\tau].$$

Note that $\mathrm{lcm}(a,b)$ is defined uniquely up to elements of $\mathbb{Z}[\tau]^\times$.

The $\mathbb{I}$-content of $q \in \mathbb{I}$ is defined as

$$\mathrm{cont}_{\mathbb{I}}(q) \;:=\; \mathrm{lcm}\{\alpha \in \mathbb{Z}[\tau]^\bullet \mid q \in \alpha\mathbb{I}\} \tag{1.32}$$

and analogously the L-content of a sublattice Λ of L as

$$\mathrm{cont}_{L}(\Lambda) \;:=\; \mathrm{lcm}\{m \in \mathbb{N} \mid \Lambda \subset mL\}. \tag{1.33}$$

LEMMA 1.18. *An icosian $q \in \mathbb{I}$ is primitive if and only if*

$$\mathrm{cont}_{\mathbb{I}}(q) \in \mathbb{Z}[\tau]^{\times}.$$

PROOF. Let $\alpha \in \mathbb{Z}[\tau]^{\bullet}$ with $\frac{1}{\alpha}q \in \mathbb{I}$. If $q \in \mathbb{I}$ is primitive this implies that $\alpha \in \mathbb{Z}[\tau]^{\times}$ and hence $\mathrm{cont}_{\mathbb{I}}(q) \in \mathbb{Z}[\tau]^{\times}$. Conversely, let $\alpha \in K$ with $\alpha q \in \mathbb{I}$ and $\alpha = \frac{\beta}{\gamma}$ where $\beta, \gamma \in \mathbb{Z}[\tau]$ are relatively prime. Note that also $\frac{1}{\gamma}q \in \mathbb{I}$ and thus we know that γ divides $\mathrm{cont}_{\mathbb{I}}(q) \in \mathbb{Z}[\tau]^{\times}$ which means that $\gamma \in \mathbb{Z}[\tau]^{\times}$ and hence $\alpha \in \mathbb{Z}[\tau]$. □

LEMMA 1.19. *A sublattice Λ of L is L-primitive if and only if*

$$\mathrm{cont}_{L}(\Lambda) = 1.$$

PROOF. This is proved analogously to Lemma 1.18. □

Clearly, for all $q \in \mathbb{I}$ the icosian $\frac{1}{\mathrm{cont}_{\mathbb{I}}(q)}q$ is $\mathbb{I}$-primitive and for every sublattice $\Lambda \subset L$ the sublattice $\frac{1}{\mathrm{cont}_{L}(\Lambda)}\Lambda$ is L-primitive. Consequently, every icosian q can be written as $q = \alpha p$ where $\alpha \in \mathbb{Z}[\tau]$ and $p \in \mathbb{I}$ is primitive. Similarly, every sublattice $\Lambda \subset L$ can be decomposed in $\Lambda = m\Gamma$ where $m \in \mathbb{N}$ and $\Gamma \subset L$ is L-primitive. Note that $\Lambda \subset L$ is L-primitive if and only if

$$\gcd\{(z_{ij}) \mid 1 \leq i, j \leq d\} = 1$$

for an integer matrix $Z = (z_{ij})$ with $B_{\Lambda} = B_L Z$.

LEMMA 1.20. *Let $\alpha\mathbb{I} \subset \mathbb{I}$ be a two-sided ideal which contains a primitive icosian p. Then, $\alpha \in \mathbb{Z}[\tau]^{\times}$, i.e. $\alpha\mathbb{I} = \mathbb{I}$.*

PROOF. By Lemma 1.10 we know that every two-sided ideal of $\mathbb{I}$ has the form $\alpha\mathbb{I}$ with $\alpha \in \mathbb{Z}[\tau]$. If $p \in \alpha\mathbb{I}$ there is an icosian $q \in \mathbb{I}$ such that $p = \alpha q$. Therefore $\frac{1}{\alpha}p = q \in \mathbb{I}$ and the primitivity of p implies that $\alpha \in \mathbb{Z}[\tau]^{\times}$. □

1.3.4. Factorisation. An icosian $p \in \mathbb{I}$, which is not a unit, is called *prime*, if $p = ab$ for $a, b \in \mathbb{I}$ implies that $a \in \mathbb{I}^{\times}$ or $b \in \mathbb{I}^{\times}$. Obviously, all associates of a prime icosian are prime, too. Moreover, it is clear that

every prime icosian is primitive. A direct consequence of Corollary 1.12 is that primes in $\mathbb{Z}[\tau]$, considered as elements of $\mathbb{I}$, cannot be prime in $\mathbb{I}$. Furthermore, we have the following characterisation of primes in $\mathbb{I}$.

LEMMA 1.21. *An icosian p is prime in $\mathbb{I}$ if and only if $|p|^2$ is prime in $\mathbb{Z}[\tau]$.*

PROOF. If p is not prime in $\mathbb{I}$ it is immediately clear that $|p|^2$ cannot be a prime in $\mathbb{Z}[\tau]$. Conversely, if $|p|^2$ is not a prime in $\mathbb{Z}[\tau]$, for example $|p|^2 = \alpha\beta$, where α and β are primes in $\mathbb{Z}[\tau]$, then we know by Lemma 1.11 that there are $a, b \in \mathbb{I}$ such that $|a|^2 = \alpha$, $|b|^2 = \beta$ and $p = ab$, hence p is not prime in $\mathbb{I}$. □

The following definition and theorem are an adaption of [**38**, Section 20.8]. We say that $a, b \in \mathbb{I}$ have a *greatest left (right) common divisor*

$$d = \operatorname{glcd}(a, b) \qquad (d = \operatorname{grcd}(a, b)),$$

if (i) d is a left (right) divisor of a and b and (ii) every common left (right) divisor of a and b is a left (right) divisor of d.

THEOREM 1.22. *For any two icosians $a, b \in \mathbb{I}^\bullet$ there is a $\operatorname{glcd}(a, b)$ and a $\operatorname{grcd}(a, b)$ in $\mathbb{I}$. They are uniquely defined up to multiplication by a unit factor from the right or left, respectively, as the generators of the ideals*

$$\operatorname{glcd}(a, b)\mathbb{I} = a\mathbb{I} + b\mathbb{I} \quad \textit{and} \quad \mathbb{I}\operatorname{grcd}(a, b) = \mathbb{I}a + \mathbb{I}b.$$

PROOF. The set $S = a\mathbb{I} + b\mathbb{I}$ is clearly a right ideal in $\mathbb{I}$. Since every right ideal in $\mathbb{I}$ is principal there is an icosian d such that $S = d\,\mathbb{I}$. As $d \in S$, there are $q_\ell, p_\ell \in \mathbb{I}$, such that $d = aq_\ell + bp_\ell$. Because $a, b \in S$, d is a common left-hand divisor of a and b. In fact, any common left-hand divisor of a and b is a left-hand divisor of any element of S including d. Hence $d = \operatorname{glcd}(a, b)$ and as generator of the right ideal S, the icosian $\operatorname{glcd}(a, b)$ is uniquely defined up to a right-hand unit factor.

The claim about $\operatorname{grcd}(a, b)$ is proved analogously. □

THEOREM 1.23. *Let $p \in \mathbb{I}$ be primitive and let $|p|^2 = \pi_1 \dots \pi_n$ be the prime factorisation of $|p|^2$ in $\mathbb{Z}[\tau]$, with an arbitrary but fixed order of factors. Then, there are prime icosians $p_1 \dots p_n$, which are unique up to elements of $\mathbb{I}^\times$, such that*

$$p = p_1 \dots p_n \qquad \text{and} \qquad |p_i|^2 = \pi_i.$$

PROOF. Let $p_1 = \operatorname{glcd}(\pi_1, p)$, i.e. let there be $r_1, s_1 \in \mathbb{I}$ such that $p = p_1 r_1$ and $\pi_1 = p_1 s_1$. Hence $\pi_1^2 = |p_1|^2 |s_1|^2$ and since π_1 is a prime in $\mathbb{Z}[\tau]$ we infer that $|p_1|^2$ is an associate of $1, \pi_1$ or π_1^2 in $\mathbb{Z}[\tau]$.

By Theorem 1.22 we know that there are $u, v \in \mathbb{I}$ such that $p_1 = \pi_1 u + pv$. Hence $|p_1|^2 = \pi_1^2 |u|^2 + |p|^2 |v|^2 + \pi_1 \operatorname{tr}(pv\overline{u})$, which reveals that π_1 divides $|p_1|^2$. With (1.22) we conclude that $|p_1|^2$ cannot be an associate of 1. Suppose that $|p_1|^2$ is an associate of π_1^2. This implies that $|s_1|^2 \in \mathbb{Z}[\tau]^\times$ and hence $s_1 \in \mathbb{I}^\times$, making π_1 and p_1 associates, which is impossible since p_1 is primitive as a divisor of p. So we have $|p_1|^2 = \pi_1$ and $p = p_1 r_1$ where $|r_1|^2 = \pi_2, \dots, \pi_n$. Moreover, by Lemma 1.21 we know that p_1 is a prime icosian.

Repeating the argument with r_1, we produce iteratively the factorisations

$$r_1 = p_2 r_2,\ r_2 = p_3 r_3,\ \dots,\ r_{n-1} = p_{n-1} r_{n-1},\ r_n = p_n$$

where $|p_i|^2 = \pi_i$. This factorisation is unique up to elements of $\mathbb{I}^\times$, since the greatest left common divisor of two icosians is by Theorem 1.22 well-defined up to a right unit factor. □

COROLLARY 1.24. *Let $\alpha, \beta \in \mathbb{Z}[\tau]$ such that $\gcd(\alpha, \beta) = 1$ in $\mathbb{Z}[\tau]$. Then, we have $\operatorname{glcd}(\alpha, \beta) = 1 = \operatorname{grcd}(\alpha, \beta)$ in $\mathbb{I}$.*

PROOF. Let $g_\ell = \operatorname{glcd}(\alpha, \beta)$ and $g_r = \operatorname{grcd}(\alpha, \beta)$. There are $u_r, u_\ell, v_r, v_\ell \in \mathbb{I}$, such that $\alpha = u_r g_r = g_\ell u_\ell$ and $\beta = v_r g_r = g_\ell v_\ell$. We see that $|g_r|^2$ and $|g_\ell|^2$ divide $\gcd(|\alpha|^2, |\beta|^2) = \gcd(\alpha^2, \beta^2) = 1$. Hence $|g_r|^2, |g_\ell|^2 \in \mathbb{Z}[\tau]^\times$ and $g_r, g_\ell \in \mathbb{I}^\times$, compare (1.22). □

COROLLARY 1.25. *If $p \in \mathbb{I}$ is primitive and $\alpha \in \mathbb{Z}[\tau]$, then*

$$|\operatorname{glcd}(p,\alpha)|^2 = \gcd(|p|^2,\alpha) = |\operatorname{grcd}(p,\alpha)|^2.$$

PROOF. Denote $\gamma = \gcd(|p|^2,\alpha)$ and let $|p|^2 = \pi_1 \dots \pi_m$ be the prime factorisation of $|p|^2$ in $\mathbb{Z}[\tau]$, such that $\gamma = \pi_1 \dots \pi_\ell$ for some $\ell \leq m$. By Theorem 1.23 there are prime icosians $p_1, \dots, p_m \in \mathbb{I}$ such that $p = p_1 \dots p_m$ and $|p_i|^2 = \pi_i$ for $i = 1, \dots, m$, i.e. $\gamma = p_1 \dots p_\ell \overline{p}_\ell \dots \overline{p}_1$. Since p is primitive, $p_1 \dots p_\ell \overline{p}_\ell$ cannot be a divisor of p, as this would mean that $|p_\ell|^2 \in \mathbb{Z}[\tau]$ divides p. Hence, $\operatorname{glcd}(p,\gamma) = p_1 \dots p_\ell$ and $|\operatorname{glcd}(p,\gamma)|^2 = \gamma$. Since $\operatorname{glcd}(p,\alpha)$ divides γ and p by definition, we infer that $\operatorname{glcd}(p,\alpha)$ divides $\operatorname{glcd}(\gamma,p)$. Conversely, $\operatorname{glcd}(\gamma,p)$ divides γ and thus α as well as p, such that $\operatorname{glcd}(\gamma,p) = \operatorname{glcd}(p,\alpha)$. Hence $|\operatorname{glcd}(p,\alpha)|^2 = \gamma = \gcd(|p|^2,\alpha)$.

Let $|p|^2 = \varphi_1 \dots \varphi_m$ be the prime factorisation of $|p|^2$ in $\mathbb{Z}[\tau]$, such that

$$\gamma = \varphi_{m-\ell+1} \cdots \varphi_m.$$

Of course, $\varphi_{m-\ell+1} \dots \varphi_m = \pi_1 \dots \pi_\ell$. By Theorem 1.23 there are prime icosians $q_1, \dots, q_m \in \mathbb{I}$, which do not necessarily coincide with $p_1, \dots p_m$, such that $p = q_1 \dots q_m$ and $|q_i|^2 = \varphi_i$ for $i = 1, \dots, m$. Hence $\gamma = \overline{q}_m \dots \overline{q}_{m-\ell+1} q_{m-\ell+1} \dots q_m$. Since p is primitive, $\overline{q}_{m-\ell+1} q_{m-\ell+1} \dots q_m$ cannot be a divisor of p, as this would mean that $|q_{m-\ell+1}|^2 \in \mathbb{Z}[\tau]$ divides p. This implies that, $\operatorname{grcd}(p,\gamma) = q_{m-\ell+1} \dots q_m$ and $|\operatorname{grcd}(p,\gamma)|^2 = \gamma$. Since $\operatorname{grcd}(p,\alpha)$ divides γ and p by definition, we infer that $\operatorname{grcd}(p,\alpha)$ divides $\operatorname{grcd}(\gamma,p)$. Conversely, $\operatorname{grcd}(\gamma,p)$ divides γ and thus α as well as p, such that $\operatorname{grcd}(\gamma,p) = \operatorname{grcd}(p,\alpha)$. Hence $|\operatorname{grcd}(p,\alpha)|^2 = \gamma = \gcd(|p|^2,\alpha)$. □

For later use we state the following lemmas.

LEMMA 1.26. *Let $a, b \in \mathbb{I}$ and $\gamma, \delta \in \mathbb{Z}[\tau]$, such that $\gcd(\gamma,\delta) = 1$. We have*

$$(\operatorname{glcd}(a,\gamma)\mathbb{I}) \cap (\operatorname{glcd}(a,\delta)\mathbb{I}) = \operatorname{glcd}(a,\gamma\delta)\mathbb{I}.$$

PROOF. With Theorem 1.22, we see immediately that

$$\operatorname{glcd}(a,\gamma\delta)\mathbb{I} = a\mathbb{I}+\gamma\delta\mathbb{I} \subset (a\mathbb{I}+\gamma\mathbb{I})\cap(a\mathbb{I}+\delta\mathbb{I}) = (\operatorname{glcd}(a,\gamma)\mathbb{I})\cap(\operatorname{glcd}(a,\delta)\mathbb{I}).$$

Conversely, let $q \in (a\mathbb{I}+\gamma\mathbb{I})\cap(a\mathbb{I}+\delta\mathbb{I})$. So there are $q_i \in \mathbb{I}$ for $1 \le i \le 6$, such that

$$q = \operatorname{glcd}(a,\gamma)q_1 = aq_2+\delta q_3 = aq_4+\gamma q_5 = \operatorname{glcd}(a,\delta)q_6$$

Here, we see that $\operatorname{glcd}(a,\gamma)$ is a left divisor of q_3, since it does not divide δ, as $\operatorname{glcd}(\gamma,\delta) = \gcd(\gamma,\delta) = 1$, and it is a left-divisor of a by definition. So there are icosians $a_\gamma, \hat{q}_3, q_7 \in \mathbb{I}$ such that $a = \operatorname{glcd}(a,\gamma)a_\gamma, q_3 = \operatorname{glcd}(a,\gamma)\hat{q}_3$ and $(a_\gamma q_2+\delta\hat{q}_3) = \operatorname{glcd}(a_\gamma,\delta)q_7$. This implies that

$$\begin{aligned} q &= \operatorname{glcd}(a,\gamma)a_\gamma q_2+\delta\operatorname{glcd}(a,\gamma)\hat{q}_3 = \operatorname{glcd}(a,\gamma)(a_\gamma q_2+\delta\hat{q}_3)\\ &= \operatorname{glcd}(a,\gamma)\operatorname{glcd}(a_\gamma,\delta)q_7 = \operatorname{glcd}(a,\gamma\delta)q_7, \end{aligned}$$

which means that $q \in \operatorname{glcd}(a,\gamma\delta)\mathbb{I}$. □

LEMMA 1.27. *For $p,q \in \mathbb{I}$ and $g := \operatorname{glcd}(p,q)$, we have $\widetilde{g} = \operatorname{grcd}(\widetilde{p},\widetilde{q})$.*

PROOF. Obviously, $p = gr$ and $q = gs$ for $r,s \in \mathbb{I}$. Hence $\widetilde{p} = \widetilde{r}\,\widetilde{g}$ and $\widetilde{q} = \widetilde{s}\,\widetilde{g}$, so $\widetilde{g}$ divides $\operatorname{grcd}(\widetilde{p},\widetilde{q}) =: h$, i.e. we can write $h = t\widetilde{g}$, where $t \in \mathbb{I}$. Now, we see that $p = \widetilde{\widetilde{p}} = \widetilde{uh} = \widetilde{h}\widetilde{u}$ and $q = \widetilde{\widetilde{q}} = \widetilde{vh} = \widetilde{h}\widetilde{v}$, where $u,v \in \mathbb{I}$, which implies that $\widetilde{h} = g\widetilde{t}$ divides g and hence $\widetilde{h} = g$. □

CHAPTER 2

Similar Sublattices

In this chapter the problem of counting similar sublattices of a general lattice $\Gamma \subset \mathbb{R}^d$ is introduced. For the root lattice A_4 this problem is solved by constructing an index preserving bijection between its primitive similar sublattices and the primitive right ideals of the icosian ring. The result is expressed in terms of a Dirichlet series generating function, which gives the number of different similar sublattices of each index. We conclude with a review of the corresponding results for the root lattices A_d with $d \leq 3$. The content of this chapter was published in [**9**].

2.1. Generalities

2.1.1. Similarities. A *similarity* σ is a non-zero linear map of $\mathbb{R}^d$ such that

$$\langle \sigma(u), \sigma(v) \rangle = c\langle u, v \rangle,$$

for a constant $c \in \mathbb{R}_+$ and $u, v \in \mathbb{R}^d$. An alternative characterisation is the following.

LEMMA 2.1. *A linear map σ of $\mathbb{R}^d$ is a similarity if and only if it is of the form $\sigma = \alpha R$ for $\alpha \in \mathbb{R}^\bullet$ and $R \in \mathrm{O}(d)$.*

PROOF. Let σ be a linear map of $\mathbb{R}^d$. If for a constant $c \in \mathbb{R}_+$ and all $u, v \in \mathbb{R}^d$ $\langle \sigma(u), \sigma(v) \rangle = c\langle u, v \rangle$, this is equivalent to $\langle \frac{1}{\sqrt{c}}\sigma(u), \frac{1}{\sqrt{c}}\sigma(v) \rangle = \langle u, v \rangle$. This is again equivalent to $\frac{1}{\sqrt{c}}\sigma \in \mathrm{O}(d)$, which means that there exits a matrix $R \in \mathrm{O}(d)$ such that $\sigma(u) = \sqrt{c}Ru$ for all $u \in \mathbb{R}^d$. As $c > 0$ we set $\sqrt{c} = \alpha \in \mathbb{R}$ and we are done. □

The set of isometries that occur in similarities that map the lattice $\Gamma \subset \mathbb{R}^d$ into itself is defined as

$$\mathrm{OS}(\Gamma) := \{R \in \mathrm{O}(d) \mid \alpha R\Gamma \subset \Gamma \text{ for some } \alpha \in \mathbb{R}_+\}\,. \tag{2.1}$$

Moreover, we define $\mathrm{SOS}(\Gamma)$ as the subset of rotations in $\mathrm{OS}(\Gamma)$.

LEMMA 2.2. *If $\Gamma \subset \mathbb{R}^d$ is a lattice, the sets $\mathrm{OS}(\Gamma)$ and $\mathrm{SOS}(\Gamma)$ are subgroups of $\mathrm{O}(d)$.*

PROOF. Obviously it is sufficient to show that $\mathrm{OS}(\Gamma)$ is a subgroup of $\mathrm{O}(d)$. If $R_1, R_2 \in \mathrm{OS}(\Gamma)$, with $\alpha_i \in \mathbb{R}$ such that $\alpha_i R_i \Gamma \subset \Gamma$ for $i = 1, 2$, then $\alpha_1\alpha_2 R_1 R_2 \Gamma = \alpha_1 R_1(\alpha_2 R_2 \Gamma) \subset \alpha_1 R_1 \Gamma \subset \Gamma$ and so $R_1 R_2 \in \mathrm{OS}(\Gamma)$.

Moreover, $\alpha R\Gamma \subset \Gamma$, with $R \in \mathrm{OS}(\Gamma)$ and $\alpha \in \mathbb{R}$, implies that $\Gamma \subset \frac{1}{\alpha} R^{-1}\Gamma$. With Lemma 1.1 we see that there is an integer matrix Z such that

$$|\det(Z)| = [\tfrac{1}{\alpha} R^{-1}\Gamma : \Gamma] = [\Gamma : \alpha R\Gamma] =: m.$$

By Lemma 1.2 , this means that $\frac{m}{\alpha} R^{-1}\Gamma \subset \Gamma$, which shows that also R^{-1} is an element of $\mathrm{OS}(\Gamma)$. □

LEMMA 2.3. *Let Γ_2 be a sublattice of $\Gamma_1 \subset \mathbb{R}^d$. Then, $\mathrm{OS}(\Gamma_1) = \mathrm{OS}(\Gamma_2)$.*

PROOF. Let $m := [\Gamma_1 : \Gamma_2]$, $\alpha \in \mathbb{R}$, and $R \in \mathrm{OS}(\Gamma_1)$ such that $\alpha R\Gamma_1 \subset \Gamma_1$. With Lemma 1.2 this implies that $m\alpha R\Gamma_2 \subset m\alpha R\Gamma_1 \subset m\Gamma_1 \subset \Gamma_2$ and hence $R \in \mathrm{OS}(\Gamma_2)$.

Conversely, we conclude with $\alpha \in \mathbb{R}_+$ and $R \in \mathrm{OS}(\Gamma_2)$ such that $\alpha R\Gamma_2 \subset \Gamma_2$, that $m\alpha R\Gamma_1 \subset \alpha R\Gamma_2 \subset \Gamma_2 \subset \Gamma_1$ and so $R \in \mathrm{OS}(\Gamma_1)$. □

Two lattices in Γ and Λ in $\mathbb{R}^d$ are called *similar* if there is a similarity σ, such that $\sigma(\Gamma) = \Lambda$. A straightforward calculation with the corresponding basis matrices gives the following

LEMMA 2.4. *Two lattices Γ and Λ in $\mathbb{R}^d$ are similar if and only if their Gram matrices satisfy the relation*

$$G_\Lambda = \alpha Z G_\Gamma Z^t,$$

where $\alpha \in \mathbb{R}^+$ and Z is an invertible integer matrix. □

2.1.2. Similar Sublattices. A sublattice $\Lambda \subset \Gamma$ is referred to as a *similar sublattice (SSL)* of Γ, if there is a similarity σ of Γ such that $\Lambda = \sigma(\Gamma)$. For a given lattice $\Gamma \subset \mathbb{R}^d$ one is interested in its *similar sublattices (SSLs)*, the possible indices of its SSLs and the number of distinct SSLs of each index.

We define the arithmetic function $f_\Gamma(m)$ as the number of distinct SSLs of index m. Clearly, $f_\Gamma(1) = 1$. Note that, f_Γ is not always multiplicative. For example in [**14**] we find several planar lattices with non-multiplicative arithmetic function f_Γ. However, in many relevant cases, and in all cases that will appear below, $f_\Gamma(m)$ is multiplicative, which implies that the corresponding Dirichlet series generating function,

$$D_\Gamma(s) \;:=\; \sum_{m=1}^{\infty} \frac{f_\Gamma(m)}{m^s}\,, \tag{2.2}$$

possesses an Euler product expansion, compare [**2**, Ch.11].

Every lattice $\Gamma \subset \mathbb{R}^d$ possesses *trivial SSLs*, i.e. SSLs of the form $m\Gamma$ where $m \in \mathbb{N}$. Note that $m\Gamma$ has index m^d in Γ. If there are no non-trivial SSLs, the generating function thus simply reads

$$D_\Gamma(s) \;=\; \zeta(ds)\,, \tag{2.3}$$

where $\zeta(s)$ is Riemann's zeta function [**2**, Ch.11]. If there are non-trivial SSLs, each of them can again be scaled by an arbitrary natural number, so that

$$D_\Gamma(s) \;=\; \zeta(ds)\, D^{\mathsf{pr}}_\Gamma(s)\,, \tag{2.4}$$

where $D_\Gamma^{\mathsf{pr}}(s)$ is the Dirichlet series generating function for the *primitive* SSLs of Γ. It is the aim of this chapter to derive this Dirichlet series for the root lattice A_4.

Let us first summarise some general properties of the generating functions in this context. A simple conjugation argument based on Lemma 2.4, see for example [**39, 24**], shows that if Γ and Λ are similar lattices in $\mathbb{R}^d$, one has $D_\Gamma(s) = D_\Lambda(s)$. Moreover, we have the following.

THEOREM 2.5. *If Γ is a lattice in $\mathbb{R}^d$, one has $D_\Gamma(s) = D_{\Gamma^*}(s)$.*

PROOF. If σ is any similarity in $\mathbb{R}^d$, one has

$$\sigma\Gamma \subset \Gamma \quad \Longleftrightarrow \quad \sigma^t\Gamma^* \subset \Gamma^*,$$

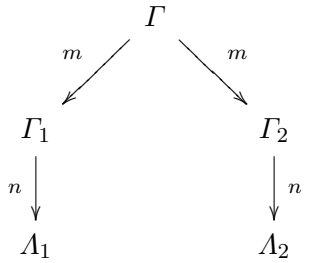
which, in view of (1.2), is immediate from the relation $\langle x|\sigma y\rangle = \langle \sigma^t x|y\rangle$; see for example [**53**, p. 524]. Observing that $\det(\sigma) = \det(\sigma^t)$, one thus obtains an index preserving bijection between the SSLs of Γ and those of Γ^*, whence we have $f_\Gamma(m) = f_{\Gamma^*}(m)$ for all $m \in \mathbb{N}$. This gives $D_\Gamma(s) = D_{\Gamma^*}(s)$. □

Although the arithmetic function f_Γ is not multiplicative in general, one can show that it is super-multiplicative, which we recall from [**39**].

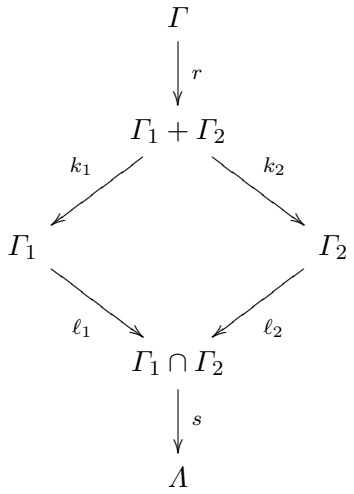
THEOREM 2.6. *The arithmetic function $f_\Gamma(m)$ of a lattice $\Gamma \subset \mathbb{R}^d$ is super-multiplicative, i.e. for coprime $m, n \in \mathbb{N}$ one has*

$$f_\Gamma(mn) \geq f_\Gamma(m) f_\Gamma(n).$$

PROOF. Let Γ_1 and Γ_2 be SSLs of Γ of index m, and for $i \in \{1, 2\}$, let Λ_i be an SSL of Γ_i of index n. This can be illustrated as follows:



where $\leftarrow$ has to be interpreted as $\subset$ and the letter next to the arrows denote the corresponding index. Since the Dirichlet series of similar lattices coincide we have $f_\Gamma(m) = f_{\Gamma_1}(m) = f_{\Gamma_2}(m)$ for all $m \in \mathbb{N}$. So we have to show that $\Lambda_1 \neq \Lambda_2$ if $\Gamma_1 \neq \Gamma_2$ or equivalently that $\Lambda_1 = \Lambda_2$ implies $\Gamma_1 = \Gamma_2$. Let $\Lambda_1 = \Lambda_2 =: \Lambda$ and consider the following diagram:



As $[\Gamma_1 : \Lambda] = [\Gamma_2 : \Lambda] = n$ we know that $\frac{n}{s} = \ell_1 = \ell_2 =: \ell$. Moreover, $\frac{m}{r} = k_1 = k_2 := k$, since $[\Gamma : \Gamma_1] = [\Gamma : \Gamma_2] = m$. The second isomorphism theorem for groups, see for example [**53**, p. 17], implies that $\Gamma_2/(\Gamma_1 \cap \Gamma_2) \simeq (\Gamma_1 + \Gamma_2)/\Gamma_1$ and hence $k = \ell$. This means that k divides m as well as n, which are coprime by assumption. This requires that $k = [\Gamma_1 : (\Gamma_1 + \Gamma_2)] = 1$ and thus $\Gamma_1 = \Gamma_2$. □

2.2. Results for A_4

2.2.1. Similarities in 4-space.

THEOREM 2.7. *All similarities in $\mathbb{R}^4 \simeq \mathbb{H}(\mathbb{R})$ can be written as either $x \mapsto pxq$ (orientation preserving case) or as $x \mapsto p\overline{x}q$ (orientation reversing case), with non-zero quaternions $p, q \in \mathbb{H}(\mathbb{R})$. The determinants of the linear maps defined this way are $\pm|p|^4|q|^4$. Conversely, all maps of the form $x \mapsto pxq$ and $x \mapsto p\overline{x}q$ are similarities.*

PROOF. By Lemma 2.1 we know that every similarity σ of $\mathbb{R}^4$ has the form $\sigma = \alpha R$ where $\alpha \in \mathbb{R}_+$ and $R \in \mathrm{O}(4)$. Hence the statement of this theorem follows from the parametrisation of $\mathrm{O}(4)$ by pairs of quaternions, see [**48, 3, 12**] for details. □

Sometimes it is convenient to refer to the standard matrix representation of the linear map $x \mapsto pxq$, which is defined via

$$M(p,q)x^t := (pxq)^t. \tag{2.5}$$

Recall that we denote quaternions as row vectors, so that x^t is a column vector. Moreover, note that on the right hand side of (2.5) we mean quaternion multiplication whereas on the left hand side we mean matrix-vector multiplication. Further details can be found in [**48, 3, 12**].

2.2.2. Similar Sublattices via Quaternions and Icosians. We want to analyse the SSLs of the root lattice A_4. As outlined in Section 1.2.1 the lattice L, from (1.5), is a scaled copy of A_4, so we can equivalently analyse the SSLs of the lattice L.

LEMMA 2.8. *All SSLs of the lattice L, as defined in* (1.5)*, are images of L under orientation preserving maps of the form $x \mapsto pxq$, with $p, q \in \mathbb{H}(K)^\bullet$ and $K = \mathbb{Q}(\sqrt{5}\,)$.*

PROOF. It is easy to check computationally that $L = \overline{L}$, i.e. L is invariant under quaternion conjugation as defined in (1.9). Hence we need not consider orientation reversing similarities. By Theorem 2.7, all SSLs are thus images of L under maps $x \mapsto pxq$ with $p, q \in \mathbb{H}(\mathbb{R})^\bullet$.

As defined in (2.5) consider the matrix $M(p,q)$, which corresponds to the map $x \mapsto pxq$. There are certain linear combinations of the matrix entries in $M(p,q)$ which show that all products $p_i q_j$ must be in the quadratic field $K = \mathbb{Q}(\sqrt{5}\,)$; see [**3**] for more details. This leaves the choice to take $p, q \in \mathbb{H}(K)^\bullet$. □

We intend to find a subset of the set of maps from Lemma 2.8 whose cardinality is as small as possible while reaching all SSLs. The first steps are provided by the following observations.

LEMMA 2.9. *If $p \in \mathbb{I}$, $pL\widetilde{p}$ is an SSL of L.*

PROOF. By Theorem 2.7, $pL\widetilde{p}$ is similar to L, so it remains to be shown that $pL\widetilde{p} \subset L$. Observe that by Lemma 1.16 $p \in \mathbb{I}$ implies $\widetilde{p} \in \mathbb{I}$, so that $pL\widetilde{p} \subset \mathbb{I}$ is clear. If x is any point of L, we have $\widetilde{x} = x$ by Proposition 1.15. Using the properties of the twist map from Lemma 1.14, one gets

$$(px\widetilde{p})^{\sim} \;=\; \widetilde{\widetilde{p}}\,\widetilde{x}\,\widetilde{p} \;=\; px\widetilde{p}\,.$$

Consequently, again by Proposition 1.15, $px\widetilde{p} \in L$, and hence $pL\widetilde{p} \subset L$, as claimed. □

PROPOSITION 2.10. *If $pLq \subset L$ with $p, q \in \mathbb{H}(K)^\bullet$, there is an $\alpha \in \mathbb{Q}$ such that*

$$q \;=\; \alpha\,\widetilde{p}\,.$$

PROOF. If $p, q \in \mathbb{H}(K)^\bullet$, the inclusion $pLq \subset L$ implies, by Proposition 1.15 and Lemma 1.14, that $pxq = \widetilde{p\widetilde{x}q} = \widetilde{q}x\widetilde{p}$ for all $x \in L$, hence also $(\widetilde{q})^{-1}px = x\widetilde{p}q^{-1} = (\widetilde{p}q^{-1})^{\sim}\,x$. Since $1 \in L$, this implies $(\widetilde{p}q^{-1})^{\sim} = \widetilde{p}q^{-1}$, and we get

$$x\widetilde{p}q^{-1} \;=\; \widetilde{p}q^{-1}x\,,$$

still for all $x \in L$. Noting that $\langle L\rangle_K = \mathbb{H}(K)$, the previous equation implies that $\widetilde{p}q^{-1}$ must be central, i.e., an element of K. Consequently, $q = \alpha\widetilde{p}$ for some $\alpha \in K$.

Since $p \in \mathbb{H}(K)$, we can choose some $0 \neq \beta \in \mathbb{Z}[\tau]$ such that $w = \beta p \in \mathbb{I}$. Observing that $(\beta p)^{\sim} = \beta'\widetilde{p}$, one sees that $\alpha\,px\widetilde{p} = \frac{\alpha}{\mathrm{N}(\beta)}\,wx\widetilde{w}$, where $0 \neq \mathrm{N}(\beta) \in \mathbb{Z}$. As $wL\widetilde{w} \subset L$ by Lemma 2.9, and since $L \cap \tau L = \{0\}$, the original relation $pLq \subset L$ now implies $\frac{\alpha}{\mathrm{N}(\beta)} \in \mathbb{Q}$, hence also $\alpha \in \mathbb{Q}$. □

LEMMA 2.11. *All SSLs of the lattice L are images of L under maps of the form $x \mapsto \alpha p x \widetilde{p}$ with $p \in \mathbb{I}$ primitive and $\alpha \in \mathbb{Q}$.*

PROOF. By Lemma 2.8 and Proposition 2.10, we know that maps of the form $x \mapsto \beta q x \widetilde{q}$ with $q \in \mathbb{H}(K)$ and $\beta \in \mathbb{Q}$ suffice to reach all SSLs of L. Since all coordinates of q are in $K = \mathbb{Q}(\sqrt{5}\,)$, there is a natural number m such that $p := mq \in \mathcal{L}$, with $\mathcal{L}$ as in (1.17). Then, with $\alpha := \beta/m^2$, one has $\alpha p x \widetilde{p} = \beta q x \widetilde{q}$ which means that the maps $x \mapsto \alpha p x \widetilde{p}$ and $x \mapsto \beta q x \widetilde{q}$ are equal. Since $\alpha \in \mathbb{Q}$, p and $\widetilde{p}$ are elements of $\mathcal{L} \subset \mathbb{I}$.

If p is primitive in $\mathbb{I}$, we are done. If not, we know that with $c = \mathrm{cont}_{\mathbb{I}}(p)$, from (1.32), p/c is a primitive element of $\mathbb{I}$. Simultaneously, we have $(p/c)^{\sim} = \widetilde{p}/c'$. Since $c \in \mathbb{Z}[\tau]$, we know that $cc' \in \mathbb{Z}$, so that this factor can be absorbed into α. □

At this stage, we recollect an important property of the icosian ring from [**12**].

THEOREM 2.12. *Let $p, q \in \mathbb{H}(K)^{\bullet}$ be such that $p\mathbb{I}q \subset \mathbb{I}$. If p or q is a primitive icosian, the other one must be in $\mathbb{I}$ as well.*

PROOF. This follows from [**12**, Proposition 1 and Remark 1], where this is shown for any maximal order of class number one. In particular, it applies to $\mathbb{I}$. □

As a result of independent interest, we note the following.

THEOREM 2.13. *The linear map ϱ defined by $x \mapsto \alpha p x \widetilde{p}$ has trace $\alpha\,\mathrm{N}(\mathrm{tr}(p))$ and determinant $\alpha^4\,\mathrm{N}(|p|^4)$, and its characteristic polynomial reads*

$$X^4 - \mathrm{trace}(\varrho)X^3 + AX^2 - BX + \det(\varrho)$$

where $A = \alpha^2\big(\mathrm{Tr}\big((\mathrm{tr}(p))^2(\mathrm{nr}(p))'\big) - 2\,\mathrm{N}(\mathrm{nr}(p))\big)$ and $B = \alpha^3\,\mathrm{N}\big(\mathrm{tr}(p)\,\mathrm{nr}(p)\big)$.

PROOF. This is a straightforward calculation with the standard matrix representation from (2.5) for the linear map ϱ, taking into account that

$\mathrm{nr}(\widetilde{p}) = (\mathrm{nr}(p))'$ and expressing the coefficients in terms of traces and norms. □

As Lemma 2.11 suggests we analyse now how an SSL of L of the form $pL\widetilde{p}$ with an $\mathbb{I}$-primitive icosian p relates to the L-primitive sublattices of L.

PROPOSITION 2.14. *If $p \in \mathbb{I}$ is $\mathbb{I}$-primitive, $pL\widetilde{p}$ is an L-primitive sublattice of the lattice L.*

PROOF. By Lemma 2.9, we know that $pL\widetilde{p} \subset L$. Thus by Lemma 1.19 we have to show that $\frac{1}{m}pL\widetilde{p} \subset L$ implies $m = 1$.

Note that $\frac{1}{m}\,pL\widetilde{p} \subset L$ implies $\frac{1}{m}\,pL[\tau]\widetilde{p} \subset L[\tau]$. From (1.28) and Lemma 1.2, we know that $5\mathbb{I} \subset L[\tau]$, which means that

$$\tfrac{5}{m}\,p\mathbb{I}\widetilde{p} \subset \tfrac{1}{m}\,pL[\tau]\widetilde{p} \subset L[\tau] \subset \mathbb{I}\,.$$

Since p is $\mathbb{I}$-primitive by assumption, then so is $\widetilde{p}$. By Theorem 2.12, this forces $5/m$ to be an element of $\mathbb{Z}[\tau]$. With $m \in \mathbb{N}$, this only leaves $m = 1$ or $m = 5$.

Observe next that $2L[\tau] \subset \mathcal{L}$. On the other hand, it is easily checked by means of the involved basis matrices that $\sqrt{5}\,\mathcal{L} \subset L[\tau]$. Together with $2\mathbb{I} \subset \mathcal{L}$, this gives

$$\tfrac{4\sqrt{5}}{m}\,p\mathbb{I}\widetilde{p} \subset \tfrac{2\sqrt{5}}{m}\,p\,\mathcal{L}\widetilde{p} \subset \tfrac{2}{m}\,pL[\tau]\widetilde{p} \subset 2L[\tau] \subset \mathcal{L} \subset \mathbb{I}\,.$$

By Theorem 2.12 again, we see that $\frac{4\sqrt{5}}{m} \in \mathbb{Z}[\tau]$, which (with $m \in \mathbb{N}$) is only possible for $m|4$. In combination with the previous restriction, this implies $m = 1$. □

The combination of Lemma 2.11 and Proposition 2.14, clarifies the relation of L-primitive SSLs and $\mathbb{I}$-primitive icosians.

COROLLARY 2.15. *The L-primitive SSLs of L are precisely the ones of the form $pL\widetilde{p}$ with p an $\mathbb{I}$-primitive icosian.*

PROOF. After Proposition 2.14, it remains to show that every primitive SSL M of L is of the form $pL\widetilde{p}$ for some primitive $p \in \mathbb{I}$. By Lemma 2.11, $M = \alpha pL\widetilde{p}$ with $\alpha \in \mathbb{Q}$ and $p \in \mathbb{I}$ primitive. By Proposition 2.14, $pL\widetilde{p}$ is already a primitive sublattice, so $\alpha = \pm 1$. □

Moreover, this leads to

COROLLARY 2.16. *Every SSL of the lattice L has the form $m\,pL\widetilde{p}$, where $m \in \mathbb{N}$ and $p \in \mathbb{I}$ is primitive.*

The next step is to find a suitable bijection that permits us to count the primitive SSLs of L of a given index.

LEMMA 2.17. *For $p \in \mathbb{I}$, one has $pL\widetilde{p} = L$ if and only if $p \in \mathbb{I}^{\times}$.*

PROOF. We have $pL\widetilde{p} \subset L$ for all $p \in \mathbb{I}$ by Lemma 2.9. An application of Lemma 1.1 and Theorem 2.13 reveals that the corresponding index is given by $|\det(M(p,\tilde{p})| = \mathrm{N}(|p|^4)$, where $M(p,\tilde{p})$ is the standard matrix representation from (2.5). Hence the equality $L = pL\widetilde{p}$ is equivalent to $\mathrm{N}\big(|p|^4\big) = 1$ which, in turn, is equivalent to $p \in \mathbb{I}^{\times}$, see (1.22). □

We need one further result to construct a bijective correspondence between primitive SSLs of L and certain right ideals of $\mathbb{I}$, which will then solve our problem.

LEMMA 2.18. *For primitive $r, s \in \mathbb{I}$, one has $r\mathbb{I} = s\mathbb{I}$ if and only if*

$$rL\widetilde{r} = sL\widetilde{s}.$$

PROOF. Since $r, s \in \mathbb{I}$, it is clear that $r\mathbb{I} = s\mathbb{I} \;\Rightarrow\; \mathbb{I} = r^{-1}s\mathbb{I} \;\Rightarrow\; r^{-1}s \in \mathbb{I}$. Similarly, $s^{-1}r \in \mathbb{I}$, so $r^{-1}s \in \mathbb{I}^{\times}$. Lemma 2.17 now implies $r^{-1}sL(r^{-1}s)^{\sim} = L$, which gives $rL\widetilde{r} = sL\widetilde{s}$.

Conversely, suppose that $rL\widetilde{r} = sL\widetilde{s}$, which gives $yL\widetilde{y} = L$ with $y := r^{-1}s$. Choose $\alpha \in K$ so that αy is a primitive element of $\mathbb{I}$, which is always possible. Then,

$$\alpha yL\widetilde{\alpha y} = \alpha\alpha' yL\widetilde{y} = \alpha\alpha' L$$

is a primitive sublattice of L by Proposition 2.14, whence $\alpha\alpha' = \pm 1$ and $\alpha y L \widetilde{\alpha y} = L$. This implies $\varepsilon := \alpha y = \alpha r^{-1} s \in \mathbb{I}^\times$ by Lemma 2.17. Then,

$$r \,=\, r\varepsilon\varepsilon^{-1} \,=\, \alpha s \varepsilon^{-1} \,\in\, \alpha\mathbb{I},$$

so that $\alpha' r \in \mathbb{I}$ due to $\alpha\alpha' = \pm 1$, where $\alpha' \in K$ by construction. Since $r \in \mathbb{I}$ is primitive as well, such a relation is only possible with $\alpha \in \mathbb{Z}[\tau]$, in view of the properties of the $\mathbb{I}$-content of r. Consequently, $\alpha\alpha' = \pm 1$ now gives $\alpha \in \mathbb{Z}[\tau]^\times$, so that y is an element of $\mathbb{I}$. Lemma 2.17 now implies $y \in \mathbb{I}^\times$, whence $y\mathbb{I} = \mathbb{I}$, and finally $s\mathbb{I} = r\mathbb{I}$. □

In view of our discussion so far, we call a right ideal of $\mathbb{I}$ *primitive* if it is of the form $p\mathbb{I}$ for some primitive $p \in \mathbb{I}$.

Proposition 2.19. *There is a bijective correspondence between the primitive right ideals of $\mathbb{I}$ and the primitive SSLs of L, defined by $p\mathbb{I} \mapsto pL\widetilde{p}$. Furthermore, one has the index formula*

$$[\mathbb{I} : p\mathbb{I}] \,=\, \mathrm{N}\big(\mathrm{nr}(p)^2\big) \,=\, \mathrm{N}\big(|p|^4\big) \,=\, [L : pL\widetilde{p}].$$

Proof. It is clear from Lemma 2.18 that the map is well-defined and injective, while Corollary 2.15 implies its surjectivity. The index relation follows from (1.25) and Theorem 2.13. □

2.2.3. Counting Similar Sublattices. Before we finally solve our original problem by deriving the Dirichlet series generating function D_L from (2.2), we recall from [**12**, Eq. (32)] that the Dirichlet series generating function for the number of non-zero primitive right ideals of $\mathbb{I}$ reads

$$\zeta_{\mathbb{I}}^{\mathsf{pr}}(s) \,=\, \frac{\zeta_{\mathbb{I}}(s)}{\zeta_K(4s)} \,=\, \frac{\zeta_K(2s)\,\zeta_K(2s-1)}{\zeta_K(4s)}. \tag{2.6}$$

Inserting the Euler product of ζ_K, see (1.16), one finds the expansion

$$\zeta_{\mathbb{I}}^{\mathsf{pr}}(s) \,=\, \frac{1+5^{-2s}}{1-5^{1-2s}} \prod_{p \equiv \pm 1 (5)} \left(\frac{1+p^{-2s}}{1-p^{1-2s}}\right)^2 \prod_{p \equiv \pm 2 (5)} \frac{1+p^{-4s}}{1-p^{2-4s}}. \tag{2.7}$$

THEOREM 2.20. *The number of SSLs of a given index is the same for the lattices L and A_4. There is an index preserving bijection between the primitive SSLs of A_4 and the primitive right ideals of $\mathbb{I}$. When $f(m)$ denotes the number of SSLs of index m^2 and $f^{\mathsf{pr}}(m)$ the number of primitive ones, the corresponding Dirichlet series generating functions read*

$$D_{A_4}(s) \;:=\; \sum_{m=1}^{\infty} \frac{f(m)}{m^{2s}} \;=\; \zeta(4s)\;\zeta_{\mathbb{I}}^{\mathsf{pr}}(s)$$

and

$$D_{A_4}^{\mathsf{pr}}(s) := \sum_{m=1}^{\infty} \frac{f^{\mathsf{pr}}(m)}{m^{2s}} \;=\; \zeta_{\mathbb{I}}^{\mathsf{pr}}(s)$$

where $\zeta(s)$ denotes Riemann's zeta function and $\zeta_{\mathbb{I}}^{\mathsf{pr}}(s)$ is defined in (2.6).

Furthermore, the possible indices are the squares of non-zero integers of the form $x^2 + xy - y^2 = \mathrm{N}(x + y\tau)$. All possible indices are realised. In particular, each possible index is also realised by a primitive SSL.

PROOF. Since the Gram matrices of L and A_4 only differ by a factor, see (1.7), Lemma 2.4 implies that the lattices are similar and thus their Dirichlet series coincide which gives the first claim.

By Proposition 2.19 there is an index preserving bijection between the primitive SSLs of L and the primitive right ideals of $\mathbb{I}$, which implies that $\zeta_{\mathbb{I}}^{\mathsf{pr}}(s) = D_L^{\mathsf{pr}}(s)$. As already mentioned in Section 2.1.2, a general SSL can be seen as an integer multiple of a primitive SSL, so we have according to (2.4) that

$$D_L(s) \;=\; \zeta(4s)\, D_L^{\mathsf{pr}}(s)\,.$$

The index characterisation either follows from [**23**] or from the index formula in Proposition 2.19. By Lemma 1.11 we know that for every $\alpha \in \mathbb{Z}[\tau]$ there is a $q \in \mathbb{I}$ such that $\alpha = \mathrm{nr}(q)$. Thus every possible index is realised. Since $f^{\mathsf{pr}}(m)$ vanishes precisely when $f(m)$ does (see below for an explicit formula), the last claim is clear. □

Inserting the Euler products of $\zeta(s)$ and $\zeta_{\mathbb{I}}^{\mathsf{pr}}(s)$, one finds the expansion of the Dirichlet series $D_{A_4}(s)$ as an Euler product,

$$D_{A_4}(s) = \tfrac{1}{(1-5^{-2s})(1-5^{1-2s})} \prod_{p\equiv\pm1\ (5)} \tfrac{1+p^{-2s}}{1-p^{-2s}} \tfrac{1}{(1-p^{1-2s})^2} \prod_{p\equiv\pm2\ (5)} \tfrac{1+p^{-4s}}{1-p^{-4s}} \tfrac{1}{1-p^{2-4s}}.$$

Consequently, the arithmetic function $f(m)$ (and also $f^{\mathsf{pr}}(m)$) is multiplicative, i.e., $f(mn) = f(m)\,f(n)$ for m, n coprime, with $f(1) = 1$. The functions f and f^{pr} are thus completely specified by their values at prime powers. The geometric series $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ with $|x| < 1$, the resulting identity $\frac{1+x}{1-x} = 1 + 2\sum_{n=1}^{\infty} x^n$ and the sums

$$\textstyle\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}, \quad \sum_{k=1}^{n} n x^{n-1} = \frac{1-(n+1)x^n+nx^{n+1}}{(1-x)^2}$$

lead together with the Cauchy product for series to the following values of f and f^{pr} at prime power p^r:

$$f(p^r) \;=\; \begin{cases} \frac{5^{r+1}-1}{4}, & \text{if } p = 5, \\ \frac{2\,(1-p^{r+1})-(r+1)(1-p^2)p^r}{(1-p)^2}, & \text{for primes } p \equiv \pm1\ (5), \\ \frac{2-p^r-p^{r+2}}{1-p^2}, & \text{for primes } p \equiv \pm2\ (5) \text{ and } r \text{ even}, \\ 0, & \text{for primes } p \equiv \pm2\ (5) \text{ and } r \text{ odd}, \end{cases}$$

(2.8)

$$f^{\mathsf{pr}}(p^r) \;=\; \begin{cases} 6\cdot 5^{r-1}, & \text{if } p = 5, \\ (r+1)p^r + 2rp^{r-1} + (r-1)p^{r-2}, & \text{for } p \equiv \pm1\ (5), \\ p^r + p^{r-2}, & \text{for } p \equiv \pm2\ (5) \text{ and } r \text{ even}, \\ 0, & \text{for } p \equiv \pm2\ (5) \text{ and } r \text{ odd}. \end{cases}$$

The first few terms of the Dirichlet series thus read

$$\begin{aligned} D_{A_4}(s) &= 1 + \tfrac{6}{4^{2s}} + \tfrac{6}{5^{2s}} + \tfrac{11}{9^{2s}} + \tfrac{24}{11^{2s}} + \tfrac{26}{16^{2s}} + \tfrac{40}{19^{2s}} + \tfrac{36}{20^{2s}} + \tfrac{31}{25^{2s}} + \tfrac{60}{29^{2s}} + \dots \\ D_{A_4}^{\mathsf{pr}}(s) &= 1 + \tfrac{5}{4^{2s}} + \tfrac{6}{5^{2s}} + \tfrac{10}{9^{2s}} + \tfrac{24}{11^{2s}} + \tfrac{20}{16^{2s}} + \tfrac{40}{19^{2s}} + \tfrac{30}{20^{2s}} + \tfrac{30}{25^{2s}} + \tfrac{60}{29^{2s}} + \dots \end{aligned}$$

where the denominators are the squares of the integers previously identified in [**23**], see also [**78**, entry A031363]. A comparison of $D_{A_4}(s)$ and $D^{\mathsf{pr}}_{A_4}(s) = \zeta^{\mathsf{pr}}_{\mathbb{I}}(s)$, reveals the origin of the various contributions. In particular, the six SSLs of index $4^2 = 16$ result from the five generators of primitive ideals of $\mathbb{I}$ of index 16 together with the SSL $2A_4$. No such extra solution exists for index $5^2 = 25$, while index $9^2 = 81$ emerges also from the SSL $3A_4$.

Due to our definition with $f(m)$ being the number of SSLs of A_4 of index m^2, the Dirichlet series generating function of the arithmetic function f is $D_{A_4}(s/2)$, which has nice analytic properties. For their derivation we need the Dirichlet character

$$\chi(n) \;:=\; \begin{cases} 0, & n \equiv 0 \;(5)\,, \\ 1, & n \equiv \pm 1 \;(5)\,, \\ -1, & n \equiv \pm 2 \;(5)\,, \end{cases} \tag{2.9}$$

see [**2**, Section 6.8] for details. Note that the corresponding L-series,

$$L(s,\chi) := \sum_{m=1}^{\infty} \chi(m) m^{-s},$$

defines an entire function on the complex plane. As $\zeta_K = \zeta(s)L(s,\chi)$, see [**87**, Chapter 11, Eq. (10)], we have

$$D_{A_4}(s/2) \;=\; \frac{\zeta_K(s)\,\zeta_K(s-1)}{L(2s,\chi)} \;=\; \frac{\zeta(s)\,\zeta(s-1)\,L(s,\chi)\,L(s-1,\chi)}{L(2s,\chi)}. \tag{2.10}$$

In particular, $D_{A_4}(s/2)$ is analytic on the half-plane $\{\sigma > 2\}$, where we write $s = \sigma + \mathrm{i}t$ as usual, due to the fact that $L(s,\chi)$ is analytic everywhere and $\zeta(s)$ is analytic except for a simple pole at $s = 1$ with residue 1, compare [**2**, Theorem 12.5]. Therefore $D_{A_4}(s/2)$ is analytic on the line $\{\sigma = 2\}$, except at $s = 2$, where we have a simple pole as the right-most singularity of $D_{A_4}(s/2)$. Consequently, one can derive the asymptotic growth of $f(m)$ from it, see [**12**, Appendix] for details. Since the value of the arithmetic function $f(m)$

fluctuates heavily, this is done via the corresponding summatory function

$$F(x) \;:=\; \sum_{m\le x} f(m) \;\sim\; \varrho\,\frac{x^2}{2}\,, \quad \text{as } x\to\infty,$$

where the growth constant is given by $\varrho \;=\; \operatorname{res}_{s=2} D_{A_4}(s/2) \;=\; \frac{\zeta_K(2)\,L(1,\chi)}{L(4,\chi)} \;=\; \frac{\zeta_K(2)\,L(1,\chi)\zeta(4)}{\zeta_K(4)}$. All the values involved here can be found in [**12**] and result in

$$\varrho \;=\; \frac{1}{2}\sqrt{5}\,\log(\tau) \;\approx\; 0.538\,011.$$

The corresponding calculations for the asymptotic behaviour of $f^{\mathsf{pr}}(m)$ are analogous and lead to

$$\varrho^{\mathsf{pr}} = \operatorname{res}_{s=2} D^{\mathsf{pr}}_{A_4}(s/2) = \frac{\zeta_K(2)\,L(1,\chi)}{\zeta_K(4)} = \frac{\zeta_K(2)\,L(1,\chi)}{\zeta(4)L(4,\chi)} = \frac{90}{\pi^4}\,\varrho \;\approx\; 0.497089.$$

2.3. Related Results

From previously published results, one can read off or easily derive the generating functions for the root lattices A_d with $d \le 3$.

THEOREM 2.21. *The Dirichlet series generating functions for the number of SSLs of the root lattices A_d with $d \le 3$ are $D_{A_1}(s) = \zeta(s)$, $D_{A_2}(s) = \zeta_{\mathbb{Q}(\xi_3)}$, with $\xi_3 = e^{2\pi i/3}$ and $\zeta_{\mathbb{Q}(\xi_3)}(s)$ the Dedekind zeta function of the cyclotomic field $\mathbb{Q}(\xi_3)$, and $D_{A_3}(s) = \zeta(3s)\,\Phi_{\mathrm{cub}}(3s)$, where*

$$\Phi_{\mathrm{cub}}(s) \;=\; \frac{1-2^{1-s}}{1+2^{-s}}\;\frac{\zeta(s)\,\zeta(s-1)}{\zeta(2s)}$$

is the generating function of the related cubic coincidence site lattice problem derived in [**3, 13**].

PROOF. A_1 is a scaled version of the integer lattice $\mathbb{Z}$, whence $D_{A_1}(s) = D_{\mathbb{Z}}(s) = \zeta(s)$. The triangular lattice A_2 is a scaled version of the ring of Eisenstein integers in $\mathbb{Q}(\xi_3)$, so that this generating function follows from [**5**, Proposition 1]. It can also be written as a product of two Dirichlet series,

$$D_{A_2}(s) \;=\; \zeta(2s)\,\frac{\zeta_{\mathbb{Q}(\xi_3)}(s)}{\zeta(2s)},$$

one for the scaling by integers and the other for the primitive SSLs.

For the primitive cubic lattice $\mathbb{Z}^3$, the SSL generating function was previously identified as $\zeta(3s)\,\Phi_{\mathrm{cub}}(3s)$ in [**11**, Theorem 5.1]. Noting that A_3 is a version of the face-centred cubic lattice, compare [**24**], one has to check that the cubic lattices share the same SSL statistics. This is a straight-forward calculation with their basis matrices and the integrality conditions for similarity transforms, similar to the one outlined in [**11**, Section 5], using the Cayley parametrisation of matrices in $\mathrm{SO}(3,\mathbb{Q})$ from [**3**], see also [**39**] for details. □

Clearly, by Theorem 2.5, one then also has the relation

$$D_{A_3}(s) \;=\; D_{A_3^*}(s) \;=\; D_{\mathbb{Z}^3}(s)\,. \tag{2.11}$$

Other completely worked out examples include the square lattice $\mathbb{Z}^2$, with $D_{\mathbb{Z}^2}(s) = \zeta_{\mathbb{Q}(i)}(s)$, see [**11**], and the cubic lattices in $d = 4$, see [**12**]. Also, various $\mathbb{Z}$-modules of rank $r > d$ in dimensions $d \leq 4$ are solved in [**11, 12, 5**]. Common to all these examples is the rather explicit use of methods from algebraic number theory or quaternions in conjunction with suitable parametrisations of the rotations involved. In higher dimensions, results are sparse, compare [**23**] and references given there, and we are not aware of any complete solution in terms of generating functions at present.

CHAPTER 3

Coincidence Site Lattices

In this chapter the problem of counting the coincidence site modules of a free $\mathbb{Z}$-module $\Gamma \subset \mathbb{R}^d$ is introduced. For the root lattice A_4 this problem is solved in several steps. First, we use the parametrisation of the similarity rotations by a primitive icosian to characterise the coincidence rotations. Then, we derive a formula for its coincidence index depending only on the icosian which defines the rotation. This requires a detour to the coincidence site modules of the icosian ring. At this point the number of coincidence rotations of the root lattice A_4 of each index could be derived. However, non-equivalent rotations may lead to the same coincidence site lattice. The argumentation up to this point was published in a summarised form in [**8, 40**]. We continue our analysis and derive conditions on the parameterising icosians when they lead to the same coincidence site lattice. Finally, the number of coincidence rotations as well as coincidence site lattices of each index is expressed in form of a Dirichlet series generating function. This is published in a summarised form in [**41**]. We conclude with a review of the corresponding results for the root lattices A_d with $d \leq 3$.

3.1. Generalities

3.1.1. Coincidence Site Modules and Lattices. Two $\mathbb{Z}$-modules (additive groups) $\Gamma, \Lambda \subset \mathbb{R}^d$ are called *commensurate*, denoted by $\Gamma \sim \Lambda$, if $\Gamma \cap \Lambda$ has finite subgroup index in Γ as well as Λ. From [**42**] we recall

LEMMA 3.1. *Let $\Gamma, \Lambda \subset \mathbb{R}^d$ be free $\mathbb{Z}$-modules of finite rank r. Then the following assertions are equivalent:*

(i) $\Gamma \sim \Lambda$.

(ii) *$\Gamma \cap \Lambda$ contains a free $\mathbb{Z}$-module of rank r.*

(iii) *$\Gamma \cap \Lambda$ is a free $\mathbb{Z}$-module of rank r.* □

Moreover, commensurateness of free $\mathbb{Z}$-modules in $\mathbb{R}^d$ of the same finite rank is an equivalence relation, see [**34, 42**].

An element $R \in \mathrm{O}(d)$ is called a *coincidence isometry* of a free $\mathbb{Z}$-module $\Gamma \subset \mathbb{R}^d$ of finite rank, if Γ and $R\Gamma$ are commensurate. The intersection $\Gamma \cap R\Gamma$ is then called the *coincidence site module (CSM)* for the isometry R. If Γ is a lattice it is also called *coincidence site lattice (CSL)* for the isometry R. The index of the submodule $\Gamma \cap R\Gamma$ in Γ is denoted by

$$\Sigma(R) \,=\, \Sigma_{\Gamma}(R) = \, [\Gamma : (\Gamma \cap R\Gamma)]\,,$$

and is referred to as *coincidence index* for the isometry R. The set of all coincidence isometries is defined as

$$\mathrm{OC}(\Gamma) \,:=\, \{R \in \mathrm{O}(d) \mid \Gamma \sim R\Gamma\} \tag{3.1}$$

and forms a subgroup of $\mathrm{O}(d)$; see [**42, 34**]. The subgroup $\mathrm{SOC}(\Gamma)$ consists of all rotations within $\mathrm{OC}(\Gamma)$. From [**42**, Example 2.7] or [**34**, Corollary 3.9] we infer that

$$\mathrm{OC}(\mathbb{I}) = \mathrm{OC}(L[\tau]) = \mathrm{OC}(\mathcal{L}) = \mathrm{O}(d, K). \tag{3.2}$$

When a a free $\mathbb{Z}$-module $\Gamma \subset \mathbb{R}^d$ of finite rank is given, the set $\Sigma(\mathrm{OC}(\Gamma))$ is called the *simple coincidence spectrum*. It may or may not possess an algebraic structure. In nice situations, $\Sigma(\mathrm{OC}(\Gamma))$ is a multiplicative monoid within $\mathbb{N}$. On top of the spectrum, one is interested in the number $g(m)$ of different CSMs of a given index m. This arithmetic function is often encapsulated into a Dirichlet series generating function,

$$\Phi_{\Gamma}(s) \,:=\, \sum_{m=1}^{\infty} \frac{g_{\Gamma}(m)}{m^s}\,, \tag{3.3}$$

which is a natural approach because it permits an Euler product decomposition when g is multiplicative. In order to derive this generating function it is often helpful to count the coincidence rotations of Γ of index m first. Let k denote the order of the rotation symmetry group of Γ, i.e. the subgroup of $\mathrm{SOC}(\Gamma)$ which consists of all rotations with coincidence index $\Sigma = 1$. Obviously, the number of coincidence rotations come in multiples of k. If $k\, g_{\mathrm{rot}}(m)$ denotes the number of coincidence rotations with coincidence index m, we define the corresponding generating function as

$$\Phi_{\Gamma}^{\mathrm{rot}}(s) := \sum_{m=1}^{\infty} \frac{g_{\Gamma}^{\mathrm{rot}}(m)}{m^s}\,. \tag{3.4}$$

Note that $0 \leq g(m) \leq g_{\mathrm{rot}}(m)$ and that $g(m) \neq 0$ if and only if $g_{\mathrm{rot}}(m) \neq 0$.

In this chapter we focus on the analysis of the CSLs of the root lattice A_4 in form of the lattice L from (1.5). The results of this analysis are encapsulated into the Dirichlet series $\Phi_L^{\mathrm{rot}}(s)$ and $\Phi_L(s)$, which turn out to possess nice Euler product expansions. Moreover, we extract their asymptotic properties for $m \to \infty$, see [**12**] and references therein for details in this context.

3.1.2. Coincidence Indices. For later use, we derive and state some factorisation properties for coincidence indices.

Lemma 3.2. *Let $\Gamma \subset \mathbb{R}^d$ be a free $\mathbb{Z}$-module of rank $r \geq d$ and let Λ be a submodule of Γ of index $[\Gamma : \Lambda] = m$. Then,* $\mathrm{OC}(\Gamma) = \mathrm{OC}(\Lambda)$ *and for any isometry R out of this group, one has*

$$\Sigma_{\Gamma} \mid m\Sigma_{\Lambda} \quad \textit{and} \quad \Sigma_{\Lambda} \mid m\Sigma_{\Gamma}. \tag{3.5}$$

Proof. If $R \in \mathrm{OC}(\Lambda)$, we know that $\Lambda \cap R\Lambda$ has finite index in Λ and hence also in Γ. As $\Lambda \cap R\Lambda \subset \Gamma \cap R\Gamma \subset \Gamma$, we infer that $\mathrm{OC}(\Lambda) \subset \mathrm{OC}(\Gamma)$. Conversely, by Lemma 1.2, we know that $m\Gamma \subset \Lambda$ and hence by the same arguments we get $\mathrm{OC}(m\Gamma) \subset \mathrm{OC}(\Lambda)$. Since $\mathrm{OC}(m\Gamma) = \mathrm{OC}(\Gamma)$ the two groups must be equal.

Let $\Gamma = \bigcup_{j=1}^{m}(t_j + \Lambda)$ be the coset decomposition of Γ with respect to Λ. Obviously, $\Gamma \cap R\Gamma = \bigcup_{j,k=1}^{m}(t_j + \Lambda) \cap (Rt_k + R\Lambda)$. If the set

$$S_{jk}(R) := (t_j + \Lambda) \cap (Rt_k + R\Lambda)$$

is not empty let $t_{jk}(R) \in S_{jk}(R)$, i.e. $t_{jk}(R) = t_j + x = Rt_k + y$ where $x \in \Lambda$ and $y \in R\Lambda$. Let $u \in \Lambda \cap R\Lambda$, hence

$$t_{jk}(R) + u = t_j + x + u = Rt_k + y + u \in S_{jk}(R),$$

which implies that $t_{jk}(R) + \Lambda \cap R\Lambda \subset S_{jk}(R)$. Let $t'_{jk}(R) = t_j + x' = Rt_k + y'$ be any element in $S_{jk}(R)$. Since $t_{jk}(R) - t'_{jk}(R) = x - x' = y - y' \in \Lambda \cap R\Lambda$, we immediately see that

$$S_{jk}(R) = t_{jk}(R) + \Lambda \cap R\Lambda.$$

Let $I \subset \{1, \ldots, m\} \times \{1, \ldots, m\}$ be the set of pairs (j, k) such that $S_{jk}(R)$ is not empty. Then

$$\Gamma \cap R\Gamma = \bigcup_{(j,k)\in I} t_{jk}(R) + \Lambda \cap R\Lambda$$

is a coset decomposition of $\Gamma \cap R\Gamma$ with respect to $\Lambda \cap R\Lambda$. Thus

$$\Sigma_\Gamma = [\Gamma : (\Gamma \cap R\Gamma)] = \frac{[\Gamma : \Lambda][\Lambda : (\Lambda \cap R\Lambda)]}{[(\Gamma \cap R\Gamma) : (\Lambda \cap R\Lambda)]} = \frac{m\Sigma_\Lambda}{|I|},$$

which proves the first claim of (3.5). Since I has a natural group structure, it is isomorphic to a subgroup of $(\Gamma/\Lambda) \times (\Gamma/\Lambda)$. By Lagrange's Theorem, see for example [**53**, Ch. 1, Proposition 2.2], $|I|$ divides m^2. Hence $m\Sigma_\Gamma = \frac{m^2\Sigma_\Lambda}{|I|}$ implies that Σ_Λ divides $m\Sigma_\Gamma$. □

Now, we return to the analysis of lattices $\Gamma \subset \mathbb{R}^d$ and recall some results from [**90**] which are needed for the derivation of the Dirichlet series.

THEOREM 3.3. *Let $\Gamma \subset \mathbb{R}^d$ be a lattice and let $R_1, R_2 \in \mathrm{OC}(\Gamma)$, such that $\Sigma(R_1)$ and $\Sigma(R_2)$ are relatively prime. Then,*

$$\Sigma(R_1R_2) = \Sigma(R_1)\Sigma(R_2).$$

□

COROLLARY 3.4. *Let $\Gamma \subset \mathbb{R}^d$ be a lattice and let $R_1, R_2 \in \mathrm{OC}(\Gamma)$, such that $\Sigma(R_1)$ and $\Sigma(R_2)$ are relatively prime. Then,*

$$\Gamma \cap R_1R_2\Gamma = \Gamma \cap R_1\Gamma \cap R_1R_2\Gamma = (\Gamma \cap R_1\Gamma) \cap R_1(\Gamma \cap R_2\Gamma).$$

□

This corollary is rather technical but it plays an important role, since it relates the CSL of R_1R_2 with some kind of multiple CSL and hence provides the basis for a decomposition of CSLs.

LEMMA 3.5. *Let Γ be a lattice and let $R_1, R_2 \in \mathrm{OC}(\Gamma)$, such that $m := \Sigma(R_1)$ and $n := \Sigma(R_2)$ are relatively prime. Then*

$$n\Gamma \cap (\Gamma \cap R_1R_2\Gamma) = n(\Gamma \cap R_1\Gamma) \qquad mR_1\Gamma \cap (\Gamma \cap R_1R_2\Gamma) = mR_1(\Gamma \cap R_2\Gamma).$$

□

This lemma is very important since it tells us how to recover the CSLs of R_1 and R_2 from the CSL of R_1R_2.

3.1.3. Similarity and Coincidence Isometries. We consider now the relation of similarity and coincidence isometries for a lattice $\Gamma \subset \mathbb{R}^d$ as far as it is needed later. A general analysis, which shows in particular that $\mathrm{SOC}(\Gamma)$ is a normal subgroup of $\mathrm{SOS}(\Gamma)$, can be found in [**35**].

LEMMA 3.6. *If R is a coincidence isometry for the lattice $\Gamma \subset \mathbb{R}^d$, there exists some $\alpha \in \mathbb{R}_+$ so that $\alpha R\Gamma \subset \Gamma$. In other words, $\mathrm{OC}(\Gamma)$ is a subgroup of $\mathrm{OS}(\Gamma)$.*

PROOF. If $R \in \mathrm{OC}(\Gamma)$, then $\Sigma(R) = [\Gamma : (\Gamma \cap R\Gamma)] = [R\Gamma : (\Gamma \cap R\Gamma)] = n \in \mathbb{N}$. By Lemma 1.2 this implies that $nR\Gamma \subset (\Gamma \cap R\Gamma) \subset \Gamma$ and we can choose $\alpha = n$. □

Define the *denominator* of a matrix $R \in \mathrm{OS}(\Gamma)$ relative to the lattice Γ as

$$\mathrm{den}_{\Gamma}(R) \,=\, \min\{\alpha \in \mathbb{R}_{+} \mid \alpha R\Gamma \subset \Gamma\}\,. \tag{3.6}$$

Clearly, as R is an isometry, one always has $\mathrm{den}_{\Gamma}(R) \geq 1$, and from $\mathrm{den}_{\Gamma}(R) R\Gamma \subset \Gamma$ one concludes that $\big(\mathrm{den}_{\Gamma}(R)\big)^{d}$ must be an integer. Consequently, $\mathrm{den}_{\Gamma}(R)$ is an algebraic integer of degree $\leq d$. Moreover, we have the following

LEMMA 3.7. *For a lattice $\Gamma \subset \mathbb{R}^d$ let $R \in \mathrm{OS}(\Gamma)$ as defined in* (2.1)*. Then,*

$$\{\alpha \in \mathbb{R}_{+} \mid \alpha R\Gamma \subset \Gamma\} \,=\, \mathrm{den}_{\Gamma}(R)\,\mathbb{N}\,.$$

PROOF. If $\alpha \in \mathbb{R}_{+}$, such that $\alpha R\Gamma \subset \Gamma$, it follows by (3.6) that $\alpha \geq \mathrm{den}_{\Gamma}(R)$. Hence we can write $\alpha = n\,\mathrm{den}_{\Gamma}(R) + r$, where $n \in \mathbb{N}$ is maximal and
$0 \leq r < \mathrm{den}_{\Gamma}(R)$. Since $\alpha R\Gamma$ and $n\,\mathrm{den}_{\Gamma}(R) R\Gamma$ are sublattices of Γ, we conclude that $rR\Gamma \subset \Gamma$ which is only possible if $r = 0$.

The converse inclusion is clear since for any $n \in \mathbb{N}$ we know that

$$n\,\mathrm{den}_{\Gamma}(R)\Gamma \subset \mathrm{den}_{\Gamma}(R)\Gamma \subset \Gamma.$$

□

LEMMA 3.8. *Let $\Gamma \subset \mathbb{R}^d$ be a lattice, with groups* $\mathrm{OS}(\Gamma)$ *and* $\mathrm{OC}(\Gamma)$ *as defined in* (2.1) *and* (3.1) *respectively. With the denominator from* (3.6)*, one has*

$$\mathrm{OC}(\Gamma) = \{R \in \mathrm{OS}(\Gamma) \mid \mathrm{den}_{\Gamma}(R) \in \mathbb{N}\}.$$

PROOF. If $\mathrm{den}_{\varGamma}(R) \in \mathbb{N}$, one has $\mathrm{den}_{\varGamma}(R)R\varGamma \subset (\varGamma \cap R\varGamma)$. Consequently, the lattices $\varGamma$ and $R\varGamma$ are commensurate, so that the inclusion

$$\{R \in \mathrm{OS}(\varGamma) \mid \mathrm{den}_{\varGamma}(R) \in \mathbb{N}\} \subset \mathrm{OC}(\varGamma)$$

is clear.

Conversely, if $R \in \mathrm{OC}(\varGamma)$, $\varGamma$ and $R\varGamma$ are commensurate by definition. In particular, one has $\varSigma(R)R\varGamma \subset \varGamma \cap R\varGamma \subset \varGamma$, so that $\varSigma(R) \in \mathrm{den}_{\varGamma}(R)\mathbb{N}$ by Lemma 3.7. As $\varSigma(R) \in \mathbb{N}$, this is only possible if $\mathrm{den}_{\varGamma}(R) \in \mathbb{Q}$. Since we also know that $\left(\mathrm{den}_{\varGamma}(R)\right)^d \in \mathbb{N}$, we conclude that $\mathrm{den}_{\varGamma}(R) \in \mathbb{N}$ and the claim follows. □

LEMMA 3.9. *Let $\varGamma \subset \mathbb{R}^d$ be a lattice. If $R_1, R_2 \in \mathrm{OC}(\varGamma)$ generate the same CSL, i.e. $\varGamma \cap R_1\varGamma = \varGamma \cap R_2\varGamma$, then $\Sigma(R_1) = \Sigma(R_2)$ and $\mathrm{den}(R_1^{-1}) = \mathrm{den}(R_2^{-1})$.*

PROOF. The statement about the coincidence indices is clear. For the statement about the denominator note that $\mathrm{den}(R_1^{-1})\Gamma \subset R_1\Gamma$. Since $\mathrm{den}(R^{-1}) \in \mathbb{N}$ it follows that $\mathrm{den}(R_1^{-1})\Gamma \subset \Gamma \cap R_1\Gamma = \Gamma \cap R_2\Gamma \subset R_2\Gamma$. Thus $\mathrm{den}(R_1^{-1})R_2^{-1}\Gamma \subset \Gamma$, which shows that $\mathrm{den}(R_1^{-1})$ is a multiple of $\mathrm{den}(R_2^{-1})$. By symmetry in R_1 and R_2, $\mathrm{den}(R_2^{-1})$ is a multiple of $\mathrm{den}(R_1^{-1})$ and so $\mathrm{den}(R_1^{-1}) = \mathrm{den}(R_2^{-1})$. □

3.2. Results for A_4

We want to start now our analysis of the CSLs of the root lattice A_4. Due to the fact that the lattice L, as defined in (1.5), is a scaled copy of the root lattice A_4 we can equivalently analyse the CSLs of L. Since $\overline{L} = L$, any orientation reversing operation can be obtained from an orientation preserving one after applying conjugation first, so we restrict ourselves to rotations only. By Lemmas 3.6 and 3.8, we know how $\mathrm{SOC}(L)$ and $\mathrm{SOS}(L)$ are related in general. Recall from Corollary 2.16 that all SSLs of L are of the form $m\,pL\widetilde{p}$ with $m \in \mathbb{N}$ and $p \in \mathbb{I}$ primitive. For a given SSL of L, now

written as $m\,pL\widetilde{p}$, the corresponding similarity rotation is then given by the map $x \mapsto \frac{1}{|p\widetilde{p}|}\,px\widetilde{p}$, which is also referred to as

$$R(p)x := \frac{1}{|p\widetilde{p}|}\,px\widetilde{p}. \tag{3.7}$$

We denote the denominator of $R(p)$ by $\mathrm{den}(p)$. Obviously, p is only unique up to a factor $\alpha \in \mathbb{Z}[\tau]^{\times}$, hence many different p result in the same rotation $R(p)$. However, all similarity rotations of L can be characterised as

$$\mathrm{SOS}(L) = \{R(p) \mid p \in \mathbb{I} \text{ is primitive } \}\,.$$

Among them we have to identify the $\mathrm{SOC}(L)$ elements, which is possible as follows.

Proposition 3.10. *Let $0 \neq q \in \mathbb{I}$ be an arbitrary icosian. Then, the lattices $\frac{1}{|q\widetilde{q}|}\,qL\widetilde{q}$ and L are commensurate if and only if $|q\widetilde{q}| \in \mathbb{N}$. If q is primitive, then* $\mathrm{den}(q) = |q\widetilde{q}|$.

Proof. If q is primitive, we know by Proposition 2.14 that $qL\widetilde{q}$ is an L-primitive sublattice of L, hence $\mathrm{den}(q) = |q\widetilde{q}|$, which is a positive integer by Lemma 3.8. When $q = \alpha p$ with $0 \neq \alpha \in \mathbb{Z}[\tau]$, one has $|q\widetilde{q}| = \mathrm{N}(\alpha)|p\widetilde{p}|$ with $\mathrm{N}(\alpha) \in \mathbb{N}$ and the claim follows from the primitive case, since $R(p) = R(q)$. □

Let us call an icosian $q \in \mathbb{I}$ *admissible* when $|q\widetilde{q}| \in \mathbb{N}$. As $\mathrm{nr}(\widetilde{q}) = \mathrm{nr}(q)'$, the admissibility of q implies that $\mathrm{N}\big(\mathrm{nr}(q)\big) = |q\widetilde{q}|^2$ is a square in $\mathbb{N}$. An immediate consequence of the previous proposition is that

$$\mathrm{SOC}(L) = \{R(p) \mid p \in \mathbb{I} \text{ is primitive and admissible } \}\,. \tag{3.8}$$

Let us summarise what all this means for the CSLs of L.

Theorem 3.11. *The CSLs of L are precisely the lattices of the form*

$$L \cap \tfrac{1}{|p\widetilde{p}|}\,pL\widetilde{p}$$

with $p \in \mathbb{I}$ primitive and admissible.

This is the first step to connect certain primitive right ideals $p\mathbb{I}$ of the icosian ring with the CSLs of L. Before we continue in this direction, note the following lemma which is needed later.

LEMMA 3.12. *Let $q \in \mathbb{I}$ be admissible and let*

$$|q|^2 = \pi_1^{r_1} \dots \pi_m^{r_m}$$

be its prime factorisation in $\mathbb{Z}[\tau]$. Then, the exponent r_i of every prime π_i which is either a ramified prime or a splitting prime occurring in the prime factorisation of $|q|^2$ but not in that of $|\widetilde{q}|^2$, is even.

PROOF. Since $|\widetilde{q}|^2 = \mathrm{nr}(q)' = (\pi_1')^{r_1} \dots (\pi_m')^{r_m}$, the condition that

$$|q\widetilde{q}| = \mathrm{nr}(q\widetilde{q})^{\frac{1}{2}} = |\pi_1\pi_1'|^{\frac{r_1}{2}} \dots |\pi_m\pi_m'|^{\frac{r_m}{2}}$$

is an integer, implies the claim. □

Let us consider now the relation of right ideals of $\mathbb{I}$ and CSLs of L.

LEMMA 3.13. *Let $r, s \in \mathbb{I}$ be primitive and admissible quaternions, with $r\mathbb{I} = s\mathbb{I}$. Then, one has $L \cap \frac{rL\tilde{r}}{|r\tilde{r}|} = L \cap \frac{sL\tilde{s}}{|s\tilde{s}|}$.*

PROOF. When $r\mathbb{I} = s\mathbb{I}$, one has $s = r\varepsilon$ for some $\varepsilon \in \mathbb{I}^\times$; see (1.22). Since, by Lemma 2.17, we then know that $\varepsilon L\widetilde{\varepsilon} = L$, one has $rL\tilde{r} = sL\tilde{s}$ in this case. As $|\varepsilon\widetilde{\varepsilon}\,|^2 = \mathrm{N}(\mathrm{nr}(\varepsilon)) = 1$, one also finds $|s\tilde{s}| = |r\tilde{r}|$. Consequently, $\frac{rL\tilde{r}}{|r\tilde{r}|} = \frac{sL\tilde{s}}{|s\tilde{s}|}$, and the CSLs of L defined by r and s are equal. □

The converse statement to Lemma 3.13 is not true, as the equality of two CSLs does *not* imply the corresponding rotations to be symmetry related. An example is provided by $r = (\tau, 2\tau, 0, 0)$ and $s = (\tau^2, \tau, \tau, 1)$, which define the same CSL, though $s^{-1}r$ is not a unit in $\mathbb{I}$. The CSL is spanned by the basis $\{(1,2,0,0), (2,-1,0,0), (\frac{3}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}), (-1, \frac{1}{2}, \frac{\tau-1}{2}, -\frac{\tau}{2})\}$. However, when

two primitive quaternions $r, s \in \mathbb{I}$ define rotations that are related by a rotation symmetry of L, one has $r\mathbb{I} = s\mathbb{I}$ as a direct consequence of Lemma 2.17.

Although the primitive elements of $\mathbb{I}$ are important in this context, we need a variant for our further discussion. Let $p \in \mathbb{I}$ be primitive and admissible. Since $\mathbb{Z}[\tau]$ is a Dedekind domain, one has the relation $(x\mathbb{Z}[\tau])^{-1} = \frac{1}{x}\mathbb{Z}[\tau]$ for any principal fractional ideal with nonzero $x \in K$, see [**52, 62**] for details. Then, the fractional ideal

$$\begin{aligned}\big(\mathrm{nr}(p)\mathbb{Z}[\tau] \cap \mathrm{nr}(\widetilde{p}\,)\mathbb{Z}[\tau]\big)^2\big(|p\widetilde{p}|^2\mathbb{Z}[\tau]\big)^{-1} \;&=\; \frac{\big(\mathrm{lcm}(\mathrm{nr}(p), \mathrm{nr}(\widetilde{p}\,))\big)^2}{|p\widetilde{p}|^2}\,\mathbb{Z}[\tau] \\ &=\; \beta_p\mathbb{Z}[\tau]\,\beta_{\tilde{p}}\mathbb{Z}[\tau]\end{aligned}$$

is a square as well, where $\beta_p := \mathrm{lcm}(\mathrm{nr}(p), \mathrm{nr}(\widetilde{p}\,))/\,\mathrm{nr}(p) \in \mathbb{Z}[\tau]$ is well-defined up to units of $\mathbb{Z}[\tau]$, as $\mathbb{Z}[\tau]$ is a principal ideal domain. Clearly, $\beta_p\mathbb{Z}[\tau]$ and $\beta_{\tilde{p}}\mathbb{Z}[\tau]$ are coprime by construction. Since their product is a square in $\mathbb{Z}[\tau]$, we have $\beta_p\mathbb{Z}[\tau] = \big(\alpha_p\mathbb{Z}[\tau]\big)^2$ for some $\alpha_p \in \mathbb{Z}[\tau]$. Explicitly, we may choose

$$\alpha_p \;=\; \sqrt{\frac{\mathrm{lcm}(\mathrm{nr}(p), \mathrm{nr}(\widetilde{p}\,))}{\mathrm{nr}(p)}} \;=\; \sqrt{\frac{\mathrm{lcm}(\mathrm{nr}(p), \mathrm{nr}(p)')}{\mathrm{nr}(p)}} \;\in\; \mathbb{Z}[\tau]\,, \tag{3.9}$$

where we assume a suitable standardisation for the lcm of two elements of $\mathbb{Z}[\tau]$. Again, α_p is only defined up to units of $\mathbb{Z}[\tau]$, therefore we implicitly work with the principal ideal $\alpha_p\mathbb{Z}[\tau]$ here. Moreover, we have the relation $\alpha_{\tilde{p}} = \widetilde{\alpha_p} = \alpha_p'$. Note that due to its definition α_p is not divisible by any inert or ramified prime of $\mathbb{Z}[\tau]$.

Let us call the icosian $\alpha_p p$ the *extension* of the primitive admissible element $p \in \mathbb{I}$, and $(\alpha_p p, \alpha_p'\widetilde{p}\,)$ the corresponding *extension pair*. In view of the form of the rotation $x \mapsto \frac{1}{|\widetilde{p}\widetilde{p}|}\,px\widetilde{p}$, it is actually rather natural to replace p and $\widetilde{p}$ by certain $\mathbb{Z}[\tau]$-multiples, $p_\alpha := \alpha_p p$ and $\widetilde{p_\alpha} = \alpha_{\tilde{p}}\,\widetilde{p}$, such that $\mathrm{nr}(p_\alpha)$ and $\mathrm{nr}(\widetilde{p_\alpha})$ have the same prime divisors in $\mathbb{Z}[\tau]$. Note that the definition of the extension pair is unique up to units of $\mathbb{Z}[\tau]$, and that one has the

relation

$$\mathrm{nr}(p_\alpha) \,=\, \mathrm{lcm}\big(\mathrm{nr}(p),\mathrm{nr}(\widetilde{p}\,)\big) \,=\, \mathrm{nr}(\widetilde{p}_\alpha) \,=\, |p_\alpha\widetilde{p}_\alpha| \,\in\, \mathbb{N}\,, \tag{3.10}$$

which will be crucial later on. The introduction of the extension pair restores some kind of symmetry of the expressions in relation to the two quaternions involved. Clearly, since the extra factors are central, this modification does not change the rotation, so that

$$\frac{px\widetilde{p}}{|p\widetilde{p}|} \,=\, \frac{p_\alpha x\widetilde{p}_\alpha}{|p_\alpha\widetilde{p}_\alpha|} \tag{3.11}$$

holds for all quaternions x.

Lemma 3.14. *For $q \in \mathbb{I}$ and $\gamma \in K$, one has $q \in \gamma\mathbb{I}$ if and only if*

$$\{\mathrm{tr}(q\overline{y}) \mid y \in \mathbb{I}\} \subset \gamma\mathbb{Z}[\tau].$$

Proof. The statement is clear for $\gamma = 0$, so assume $\gamma \neq 0$. If $q \in \mathbb{I}$, one has $q\overline{y} \in \mathbb{I}$ and hence $\mathrm{tr}(q\overline{y}) \in \mathbb{Z}[\tau]$ for all $y \in \mathbb{I}$. Thus $\frac{1}{\gamma}q \in \mathbb{I}$ implies $\frac{1}{\gamma}\,\mathrm{tr}(q\overline{y}) \in \mathbb{Z}[\tau]$. Conversely, $\mathrm{tr}(q\overline{y}) \in \mathbb{Z}[\tau]$ for all $y \in \frac{1}{\gamma}\mathbb{I}$ means $q \in \big(\frac{1}{\gamma}\mathbb{I}\big)^* = \gamma\mathbb{I}^* = \gamma\mathbb{I}$, by Lemma 1.13, which implies the claim. □

Lemma 3.15. *If $p \in \mathbb{I}$ is primitive, there is a quaternion $z \in \mathbb{I}$ with $\mathrm{tr}(p\overline{z}) = 1$. When, in addition, p is also admissible, there exists a quaternion $z \in \mathbb{I}$ such that $\mathrm{tr}(p_\alpha\,\overline{z}) + \mathrm{tr}(\widetilde{p_\alpha\overline{z}}\,) = 1$, where p_α denotes the extension of p.*

Proof. When $p \in \mathbb{I}$, one has $\gcd\{\mathrm{tr}(p\overline{x})\mathbb{Z}[\tau] \mid x \in \mathbb{I}\} = \gamma\mathbb{Z}[\tau]$ with $\gamma \in \mathbb{Z}[\tau]$. If γ was *not* a unit in $\mathbb{Z}[\tau]$, Lemma 3.14 would imply that $p \in \gamma\mathbb{I}$, which would contradict the primitivity of p. Hence γ is a unit and so $\gamma\mathbb{Z}[\tau] = \mathbb{Z}[\tau]$. Since $\mathbb{Z}[\tau]$ is noetherian, there are *finitely* many icosians $x_i \in \mathbb{I}$, say ℓ of them, such that

$$\gcd\{\mathrm{tr}(p\overline{x}_i)\mathbb{Z}[\tau] \mid 1 \leq i \leq \ell\} \,=\, \mathrm{tr}(p\overline{x}_1)\mathbb{Z}[\tau] + \ldots + \mathrm{tr}(p\overline{x}_\ell)\mathbb{Z}[\tau] \,=\, \mathbb{Z}[\tau]\,.$$

This implies the existence of numbers $\beta_i \in \mathbb{Z}[\tau]$ for $1 \leq i \leq \ell$, such that

$$\sum_{i=1}^{\ell} \beta_i \operatorname{tr}(p\overline{x}_i) = \operatorname{tr}(p\overline{z}) = 1,$$

where $z := \sum_{i=1}^{\ell} \beta_i x_i$.

For the second claim, assume that p is also admissible and denote its extension by p_α. Let $z \in \mathbb{I}$ be the icosian from the first part of the proof, so that $\operatorname{tr}(p\overline{z}) = 1$. Since $p_\alpha = \alpha_p p$ with $\alpha_p \in \mathbb{Z}[\tau]$, this implies $\operatorname{tr}(p_\alpha \overline{z}) = \alpha_p$ and thus also

$$\alpha'_p \,=\, \widetilde{\alpha_p} \,=\, \operatorname{tr}\big(\widetilde{p_\alpha \overline{z}}\,\big) \,=\, \operatorname{tr}(\widetilde{\overline{z}}\,\widetilde{p_\alpha}).$$

Since the ideals $\alpha_p\mathbb{Z}[\tau]$ and $\alpha'_p\mathbb{Z}[\tau]$ are relatively prime by construction, we have $\alpha_p\mathbb{Z}[\tau] + \alpha'_p\mathbb{Z}[\tau] = \mathbb{Z}[\tau]$ and thus the existence of $\beta, \delta \in \mathbb{Z}[\tau]$ with $\beta\alpha_p + \delta\alpha'_p = 1$. The icosians $x = \beta z$ and $y = \delta' z$ then satisfy $\operatorname{tr}(p_\alpha \overline{x}) + \operatorname{tr}(\widetilde{\overline{y}}\,\widetilde{p_\alpha}) = 1$ as well as $\operatorname{tr}(\widetilde{\overline{x}}\,\widetilde{p_\alpha}) + \operatorname{tr}(p_\alpha \overline{y}) = 1$, where the second identity follows from the first via $\big(\operatorname{tr}(u\overline{v})\big)^{\sim} = \operatorname{tr}(\widetilde{\overline{v}}\,\widetilde{u}\,)$.

Finally, observe that $\operatorname{tr}(u\overline{v}) \in K$ for all $u, v \in \mathbb{H}(K)$, so that one also has the relation $(\operatorname{tr}(u\overline{v}))' = \operatorname{tr}(\widetilde{\overline{v}}\,\widetilde{u})$. Consequently, defining $z = \tau x + (1-\tau)y$ with the x, y from above, z is an icosian that satisfies

$$\operatorname{tr}(p_\alpha \overline{z}) + \operatorname{tr}(\widetilde{\overline{z}}\,\widetilde{p_\alpha}) \,=\, \tau\big(\operatorname{tr}(p_\alpha \overline{x}) + \operatorname{tr}(\widetilde{\overline{y}}\,\widetilde{p_\alpha})\big) + (1-\tau)\big(\operatorname{tr}(p_\alpha \overline{y}) + \operatorname{tr}(\widetilde{\overline{x}}\,\widetilde{p_\alpha})\big) \,=\, 1\,,$$

which establishes the second claim. □

Our further discussion requires the definition of the following sublattice of L

$$L(q) \,=\, \{qx + \widetilde{x}\widetilde{q} \mid x \in \mathbb{I}\} \,=\, \varphi_+(q\mathbb{I}), \tag{3.12}$$

where $q \in \mathbb{I}$. This sublattice is a generalisation of the lattice in Proposition 1.15. Note that, due to $\widetilde{\mathbb{I}} = \mathbb{I}$, one has $L(q) = \widetilde{L(q)}$.

THEOREM 3.16. *Let $p \in \mathbb{I}$ be primitive and admissible, and let $p_\alpha = \alpha_p p$ be its extension. Then, the CSL defined by p is given by*

$$L \cap \frac{1}{|p\widetilde{p}|} pL\widetilde{p} \;=\; L(p_\alpha)\,,$$

with $L(p_\alpha)$ defined as in (3.12).

PROOF. To show the claim, we establish two inclusions, where we use the fact that p and p_α define the same rotation; compare (3.11).

First, since $L(p_\alpha) \subset L$ is clear, we need to show that $|p_\alpha \widetilde{p_\alpha}|\, L(p_\alpha) \subset p_\alpha L\widetilde{p_\alpha}$. If $x \in L(p_\alpha)$, there is some $y \in \mathbb{I}$ with $x = p_\alpha y + \widetilde{y}\widetilde{p_\alpha}$. Consequently, observing the norm relations from (3.10), we find

$$|p_\alpha \widetilde{p_\alpha}|\, x \;=\; p_\alpha y\, \widetilde{\overline{p}_\alpha} \widetilde{p_\alpha} + p_\alpha \overline{p_\alpha}\, \widetilde{y}\widetilde{p_\alpha} \;=\; p_\alpha\, (y\, \widetilde{\overline{p}_\alpha} + \overline{p_\alpha}\widetilde{y})\, \widetilde{p_\alpha} \;\in\; p_\alpha L(\overline{p_\alpha})\widetilde{p_\alpha} \;\subset\; p_\alpha L\widetilde{p_\alpha}\,,$$

which gives the first inclusion.

Conversely, let $x \in L \cap \frac{1}{|p\widetilde{p}|} pL\widetilde{p}$, i.e. there is some $y \in L$ so that $x = \frac{p\, y\, \widetilde{p}}{|p\widetilde{p}|} = \frac{p_\alpha y\, \widetilde{p_\alpha}}{|p_\alpha\, \widetilde{p_\alpha}|}$. Moreover, by Lemma 3.15 there exists an icosian z such that $\mathrm{tr}(p_\alpha \overline{z}) + \mathrm{tr}(\widetilde{\overline{z}}\, \widetilde{p_\alpha}) = 1$. Observing $x = \widetilde{x}$ and the norm relations in (3.10), one finds

$$\begin{aligned} x \;=\; \mathrm{tr}(p_\alpha \overline{z})x + \widetilde{x}\,\mathrm{tr}(\widetilde{\overline{z}}\,\widetilde{p}_\alpha) \;&=\; (p_\alpha \overline{z} + z\,\overline{p}_\alpha)x + \widetilde{x}(\widetilde{\overline{z}}\,\widetilde{p}_\alpha + \widetilde{\overline{p}}_\alpha \widetilde{z}\,) \\ &=\; p_\alpha(\overline{z}x + \widetilde{y}\widetilde{z}\,) + (\widetilde{x}\widetilde{\overline{z}} + zy)\widetilde{p_\alpha}\,, \end{aligned}$$

which shows that $x \in L(p_\alpha)$. □

COROLLARY 3.17. *Let $p \in \mathbb{I}$ be admissible and primitive and let $p_\alpha = \alpha_p p$ be its extension. Then, another representation of the CSL defined by p is*

$$L(p_\alpha) = (p_\alpha \mathbb{I} + \mathbb{I}\widetilde{p}_\alpha) \cap L.$$

PROOF. Clearly, $L(p_\alpha) \subset p_\alpha \mathbb{I} + \mathbb{I}\widetilde{p}_\alpha$ and $L(p_\alpha) \subset L$, i.e.

$$L(p_\alpha) \subset (p_\alpha \mathbb{I} + \mathbb{I}\widetilde{p}_\alpha) \cap L.$$

If $q = p_\alpha x + y\widetilde{p}_\alpha \in (p_\alpha \mathbb{I} + \mathbb{I}\widetilde{p}_\alpha) \cap L$, then $q = \widetilde{q}$ and

$$q = p_\alpha x + y\widetilde{p}_\alpha = \tau(p_\alpha x + y\widetilde{p}_\alpha) + (1-\tau)(p_\alpha \widetilde{y} + \widetilde{x}\widetilde{p}_\alpha) = p_\alpha z + \widetilde{z}\,\widetilde{p}_\alpha,$$

where $z = \tau x + (1-\tau)\widetilde{y} \in \mathbb{I}$. This shows that $q \in L(p_\alpha)$. □

Now we have explicitly identified the CSLs of L. For calculating their indices we need to analyse the corresponding indices of the coincidence submodules of $\mathbb{I}$ and $L[\tau]$.

3.3. Coincidence Site Modules of $L[\tau]$ and $\mathbb{I}$

LEMMA 3.18. *Let $p \in \mathbb{I}$ be primitive and admissible, then*

$$L[\tau] \cap \tfrac{1}{|p\widetilde{p}|} pL[\tau]\widetilde{p} = L(p) \oplus \tau L(p) \quad \textit{and} \quad \Sigma_{L[\tau]}(p) = \Sigma_L^2(p).$$

PROOF. Since the lattice L is rational and the lattice τL is not, the coincidence rotations of L do not mix vectors form L and τL, so we have

$$\begin{aligned} L[\tau] \cap \frac{1}{|p\widetilde{p}|} pL[\tau]\widetilde{p} &= (L \oplus \tau L) \cap (\frac{1}{|p\widetilde{p}|} pL\widetilde{p} \oplus \frac{\tau}{|p\widetilde{p}|} pL\widetilde{p}) \\ &= (L \cap \frac{1}{|p\widetilde{p}|} pL\widetilde{p}) \oplus (\tau L \cap \frac{\tau}{|p\widetilde{p}|} pL\widetilde{p}) \\ &= L(p) \oplus \tau L(p). \end{aligned}$$

A straightforward calculation with explicit cosets implies that

$$\Sigma_{L[\tau]}(p) = [L \oplus \tau L : L(p) \oplus \tau L(p)] = [L : L(p)]^2 = \Sigma_L^2(p).$$

□

With (3.2) this implies that

$$\mathrm{SOC}(L) \subset \mathrm{OC}(L[\tau]) = \mathrm{OC}(\mathbb{I}). \tag{3.13}$$

Therefore, we continue our analysis of $\mathrm{SOC}(L)$ by analysing $\mathrm{OC}(\mathbb{I})$.

Let $(p,q) \in \mathbb{I} \times \mathbb{I}$. The corresponding rotation $x \mapsto \frac{1}{|pq|}pxq$ is also referred to as

$$R(p,q)x := \frac{1}{|pq|}pxq. \tag{3.14}$$

A pair $(p,q) \in \mathbb{I}\times\mathbb{I}$ is called *primitive*, if p and q are primitive, and it is called *admissible*, if $|pq| \in \mathbb{Z}[\tau]$. For a primitive admissible pair $(p,q) \in \mathbb{I} \times \mathbb{I}$ we obviously have that $|pq|R(p,q)\mathbb{I} = p\mathbb{I}q \subset \mathbb{I} \cap R(p,q)\mathbb{I}$, and hence Lemma 3.1 implies that $R(p,q)$ is a coincidence isometry of $\mathbb{I}$. We denote its coincidence index by $\Sigma_{\mathbb{I}}(p,q)$. The *denominator* of

$$R \in \mathrm{OS}(\mathbb{I}) := \{R \in \mathrm{O}(4) \mid \alpha R\mathbb{I} \subset \mathbb{I} \text{ for some } \alpha \in \mathbb{R}_+\}$$

is defined as $\mathrm{den}_{\mathbb{I}}(R) \in \mathbb{R}_+$, such that $[\mathbb{I} : \mathrm{den}_{\mathbb{I}}(R)R\mathbb{I}] \leq [\mathbb{I} : \alpha R\mathbb{I}]$ for all $\alpha \in \mathbb{R}_+$ with $\alpha R\mathbb{I} \subset \mathbb{I}$. Note that analogously to Lemma 3.7 we have

$$\{\alpha \in \mathbb{R}_+ \mid \alpha R\mathbb{I} \subset \mathbb{I}\} = \mathrm{den}_{\mathbb{I}}(R)\mathbb{Z}[\tau]$$

and hence $\mathrm{den}_{\mathbb{I}}(R)$ is defined uniquely up to units of $\mathbb{Z}[\tau]$. For a primitive and admissible pair $(p,q) \in \mathbb{I} \times \mathbb{I}$ we obviously have

$$\mathrm{den}_{\mathbb{I}}(R(p,q)) = |pq|. \tag{3.15}$$

In analogy to (3.9),we define for an admissible pair of icosians $(p,q) \in \mathbb{I} \times \mathbb{I}$

$$\alpha_p := \sqrt{\frac{\mathrm{lcm}(\mathrm{nr}(p),\mathrm{nr}(q))}{\mathrm{nr}(p)}}, \qquad \alpha_q := \sqrt{\frac{\mathrm{lcm}(\mathrm{nr}(p),\mathrm{nr}(q))}{\mathrm{nr}(q)}} \in \mathbb{Z}[\tau]\,. \tag{3.16}$$

Note that α_p^2 divides $\mathrm{nr}(q)$ and α_q^2 divides $\mathrm{nr}(p)$. The extended pair is defined as $(p_\alpha, q_\alpha) := (\alpha_p p, \alpha_q q)$. By definition it is clear that

$$\mathrm{nr}(p_\alpha) = \mathrm{lcm}(\mathrm{nr}(p),\mathrm{nr}(q)) = \mathrm{nr}(q_\alpha) = |p_\alpha q_\alpha| = \alpha_p\alpha_q|pq| \in \mathbb{Z}[\tau] \tag{3.17}$$

as well as

$$\alpha_q|q|^2 = \alpha_p|pq| \quad \text{and} \quad \alpha_p|p|^2 = \alpha_q|pq|. \tag{3.18}$$

Since the additional factors are central, the pair (p, q) and its extension (p_α, q_α) define the same rotation, i.e.

$$R(p,q)x = \frac{pxq}{|pq|} = \frac{p_\alpha x q_\alpha}{|p_\alpha q_\alpha|} \tag{3.19}$$

holds for all quaternions x.

LEMMA 3.19. *Let (p, q) be a primitive admissible pair of icosians and (p_α, q_α) its extension. Then there exist icosians $u, v \in \mathbb{I}$ such that* $\operatorname{tr}(p_\alpha \overline{u}) + \operatorname{tr}(\overline{v} q_\alpha) = 1$.

PROOF. Note that $\operatorname{tr}(\overline{v}q) = \operatorname{tr}(q\overline{v})$ so that the claim can be proved analogously to Lemma 3.15. □

THEOREM 3.20. *Let (p_α, q_α) be the extension of a primitive admissible pair of icosians (p, q). Then*

$$\mathbb{I} \cap \tfrac{1}{|pq|} p\mathbb{I}q = p_\alpha \mathbb{I} + \mathbb{I} q_\alpha .$$

PROOF. By (3.19) we can equivalently show that $\mathbb{I} \cap \frac{1}{|p_\alpha q_\alpha|} p_\alpha \mathbb{I} q_\alpha = p_\alpha \mathbb{I} + \mathbb{I} q_\alpha$. To show this equality, we establish two inclusions.

Since $p_\alpha, q_\alpha \in \mathbb{I}$ it is clear that $p_\alpha \mathbb{I} + \mathbb{I} q_\alpha \subset \mathbb{I}$. So we only need to show that $|p_\alpha q_\alpha|(p_\alpha \mathbb{I} + \mathbb{I} q_\alpha) \subset p_\alpha \mathbb{I} q_\alpha$. If $x \in p_\alpha \mathbb{I} + \mathbb{I} q_\alpha$, there are some $r, s \in \mathbb{I}$ such that $x = p_\alpha r + s q_\alpha$. With (3.17) this implies that

$$|p_\alpha q_\alpha| x = p_\alpha r \operatorname{nr}(q_\alpha) + \operatorname{nr}(p_\alpha) s q_\alpha = p_\alpha r \overline{q_\alpha} q_\alpha + p_\alpha \overline{p_\alpha} s q_\alpha = p_\alpha (r \overline{q_\alpha} + \overline{p_\alpha} s) q_\alpha .$$

Due to the fact that $\overline{\mathbb{I}} = \mathbb{I}$, this means that $|p_\alpha q_\alpha| x \in p_\alpha \mathbb{I} q_\alpha$ and we have proved the first inclusion.

The converse inclusion relies on Lemma 3.19. Let $u, v \in \mathbb{I}$ be the icosians such that $\operatorname{tr}(p_\alpha \overline{u}) + \operatorname{tr}(\overline{v} q_\alpha) = 1$. If $x \in \mathbb{I} \cap \frac{1}{|p_\alpha q_\alpha|} p_\alpha \mathbb{I} q_\alpha$, there is a $y \in \mathbb{I}$ such that $x = \frac{p_\alpha y q_\alpha}{|p_\alpha q_\alpha|}$. Thus

$$\begin{aligned} x &= \operatorname{tr}(p_\alpha \overline{u}) x + x \operatorname{tr}(\overline{v} q_\alpha) \\ &= p_\alpha \overline{u} x + u \overline{p_\alpha} x + x \overline{v} q_\alpha + x \overline{q_\alpha} v \end{aligned}$$

$$= p_\alpha \overline{u}x + uyq_\alpha + x\overline{v}q_\alpha + p_\alpha yv$$
$$= p_\alpha(\overline{u}x + yv) + (uy + x\overline{v})q_\alpha,$$

which shows that $x \in p_\alpha\mathbb{I} + \mathbb{I}q_\alpha$. □

COROLLARY 3.21. *Let $p \in \mathbb{I}$ be primitive and admissible, and let $p_\alpha = \alpha_p p$ be its extension. Then*

$$\mathbb{I} \cap \frac{1}{|p\widetilde{p}|} p\mathbb{I}\widetilde{p} = p_\alpha\mathbb{I} + \mathbb{I}\widetilde{p}_\alpha.$$

COROLLARY 3.22. *Let (p_α, q_α) be the extension of a primitive admissible pair of icosians (p, q). Then*

$$\mathbb{I} \cap \tfrac{1}{|pq|} p\mathbb{I}q = p_\ell\mathbb{I} + \mathbb{I}q_r,$$

where $p_\ell = \operatorname{glcd}(p_\alpha, |pq|)$ and $q_r = \operatorname{grcd}(q_\alpha, |pq|)$.

PROOF. Obviously, $p\overline{p}\,\mathbb{I}\,\overline{q}q \subset p\mathbb{I}q$, so it is clear that $|pq|\mathbb{I} \subset \mathbb{I} \cap \frac{1}{|pq|} p\mathbb{I}q$. The application of Theorem 3.20 gives

$$\mathbb{I} \cap \tfrac{1}{|pq|} p\mathbb{I}q = \mathbb{I} \cap \tfrac{1}{|pq|} p\mathbb{I}q + |pq|\mathbb{I} = p_\alpha\mathbb{I} + \mathbb{I}q_\alpha + |pq|\mathbb{I} = p_\ell\mathbb{I} + \mathbb{I}q_r.$$

□

THEOREM 3.23. *Let $p \in \mathbb{I}$ be primitive and admissible, then*

$$\Sigma_{\mathbb{I}}(p, \widetilde{p}) = \Sigma_{L[\tau]}(p, \widetilde{p}).$$

PROOF. In this proof $p \in \mathbb{I}$ is arbitrary but fixed, so we write $\Sigma_{\mathbb{I}} = \Sigma_{\mathbb{I}}(p, \widetilde{p})$ and $\Sigma_{L[\tau]} = \Sigma_{L[\tau]}(p, \widetilde{p})$ as well as $R = R(p, \widetilde{p})$ for better readability. Recall from (1.28) that $L[\tau]$ is a submodule of index 5 in $\mathbb{I}$. Since

$$R \in \mathrm{SOC}(L) \subset \mathrm{OC}(L[\tau]) = \mathrm{OC}(\mathbb{I}),$$

compare (3.13), Lemma 3.2 implies that

$$\Sigma_{\mathbb{I}} = 5\Sigma_{L[\tau]}, \quad 5\Sigma_{\mathbb{I}} = \Sigma_{L[\tau]} \text{ or } \Sigma_{\mathbb{I}} = \Sigma_{L[\tau]}. \tag{3.20}$$

We will show that the first and the second relation are impossible, thus proving the third.

Assume that $\Sigma_{\mathbb{I}} = 5\Sigma_{L[\tau]}$. Since

$$\Sigma_{\mathbb{I}} = [\mathbb{I} : (\mathbb{I} \cap \tfrac{1}{|\widetilde{pp}|} p\mathbb{I}\widetilde{p})] = 5\Sigma_{L[\tau]} = [\mathbb{I} : L[\tau]]\,[L[\tau] : (L[\tau] \cap \tfrac{1}{|\widetilde{pp}|} pL[\tau]\widetilde{p})]$$

and $L[\tau] \cap \frac{1}{|\widetilde{pp}|} pL[\tau]\widetilde{p} \subset \mathbb{I} \cap \frac{1}{|\widetilde{pp}|} p\mathbb{I}\widetilde{p}$ the assumption is equivalent to

$$L[\tau] \cap \tfrac{1}{|\widetilde{pp}|} pL[\tau]\widetilde{p} = \mathbb{I} \cap \tfrac{1}{|\widetilde{pp}|} p\mathbb{I}\,\widetilde{p} = p_\alpha \mathbb{I} + \mathbb{I}\,\widetilde{p}_\alpha.$$

Hence $p_\alpha \mathbb{I} \subset L[\tau]$ and $\mathbb{I}\widetilde{p}_\alpha \subset L[\tau]$. By Corollary 1.17 we know that $\sqrt{5}\mathbb{I} \subset L[\tau]$. Let $q := \mathrm{glcd}(p_\alpha, \sqrt{5})$, i.e. $q\mathbb{I} \subset L[\tau]$. Since we know by Lemma 1.27 that $\widetilde{q} = \mathrm{grcd}(\widetilde{p}_\alpha, -\sqrt{5})$, we infer that $\mathbb{I}\widetilde{q} \subset L[\tau]$. Proposition 1.16 tells us that for $r \in \mathbb{I}$, we have $qr - \widetilde{r}\widetilde{q} \in \sqrt{5}\mathbb{I}$, which implies that $\widetilde{r}\widetilde{q} = q(r-s)$, for $s \in \mathbb{I}$, and hence $\mathbb{I}\widetilde{q} \subset q\mathbb{I}$. As $[\mathbb{I} : q\mathbb{I}] = \mathrm{N}(\mathrm{nr}(q)^2) = [\mathbb{I} : \mathbb{I}\widetilde{q}]$, see Lemma 1.9, it follows that $q\mathbb{I} = \mathbb{I}\widetilde{q}$. Due to its definition α_p is not divisible by $\sqrt{5}$, i.e. $\gcd(\alpha_p, \sqrt{5}) = 1$, which implies by Corollary 1.24 that $q = \mathrm{glcd}(p_\alpha, \sqrt{5}) = \mathrm{glcd}(p, \sqrt{5}) \in \mathbb{I} \setminus \mathbb{Z}[\tau]$. As $q\mathbb{I} = \mathbb{I}\widetilde{q}$ is a two-sided ideal this contradicts Lemma 1.10.

Now, suppose that $5\Sigma_{\mathbb{I}} = \Sigma_{L[\tau]}$ or equivalently $25\Sigma_{\mathbb{I}} = 5\Sigma_{L[\tau]}$. Hence

$$[(\mathbb{I} \cap R\mathbb{I}) : (L[\tau] \cap RL[\tau])] = \frac{[\mathbb{I} : L[\tau][L[\tau] : (L[\tau] \cap RL[\tau])]}{[\mathbb{I} : (\mathbb{I} \cap R\mathbb{I})]} = 25. \tag{3.21}$$

For $x \in L[\tau]$ we have

$$|Rx - \widetilde{R}x|^2 = |\frac{px\widetilde{p}}{|p\widetilde{p}|} - \frac{\widetilde{(px\widetilde{p})}}{|p\widetilde{p}|}|^2 = \tfrac{1}{|\widetilde{pp}|^2}|p(x-\widetilde{x})\widetilde{p}|^2 = |x-\widetilde{x}|^2,$$

which implies by Proposition 1.16 that $\mathbb{I} \cap \frac{1}{|\widetilde{pp}|} pL[\tau]\widetilde{p} \subset L[\tau]$ and hence

$$\mathbb{I} \cap \tfrac{1}{|\widetilde{pp}|} pL[\tau]\widetilde{p} = L[\tau] \cap \tfrac{1}{|\widetilde{pp}|} pL[\tau]\widetilde{p}.$$

If we consider the coset decomposition of $L[\tau]$ in $\mathbb{I}$, where $x_0 = y_0 = 0$ and $x_i, y_j \in \mathbb{I}$ for $1 \leq i \leq 4$, i.e.

$$\mathbb{I} \cap \tfrac{1}{|\widetilde{pp}|} p\mathbb{I}\widetilde{p} = \bigcup_{i,j=0}^{4} (x_i + L[\tau]) \cap \tfrac{1}{|\widetilde{pp}|} p(y_j + L[\tau])\widetilde{p},$$

it is clear that there cannot be 25 distinct cosets of $L[\tau] \cap RL[\tau]$ in $\mathbb{I} \cap R\mathbb{I}$, since $(x_i + L[\tau]) \cap \frac{1}{|\widetilde{pp}|} pL[\tau]\widetilde{p} = \varnothing$, if $i \neq 0$. This contradicts (3.21) and hence the second case of (3.20) is ruled out, too. □

LEMMA 3.24. *Let (p, q) be a primitive admissible pair of icosians and (p_α, q_α) its extension. Then,*

$$|\operatorname{glcd}(p, |pq|)|^2 = \tfrac{1}{\alpha_p}|pq| = \alpha_q \operatorname{gcd}(|p|^2, |q|^2) \quad \textit{and}$$
$$|\operatorname{grcd}(q, |pq|)|^2 = \tfrac{1}{\alpha_q}|pq| = \alpha_p \operatorname{gcd}(|p|^2, |q|^2).$$

PROOF. Recall that $\operatorname{gcd}(|p|^2), |q|^2) \operatorname{lcm}(|p|^2, |q|^2) = |pq|^2$. Hence considering the definitions of α_p and α_q in (3.16) it is clear that $|pq| = \alpha_p \alpha_q \operatorname{gcd}(|p|^2, |q|^2)$. So in each of the two cases the second of the two equalities is clear.

Moreover, α_p, α_q are by definition relatively prime. With Corollary 1.25 and (3.18), we conclude that

$$\begin{aligned}
|\operatorname{glcd}(p, |pq|)|^2 &= \operatorname{gcd}(|p|^2, |pq|) = \operatorname{gcd}(\alpha_q |p|^2, \alpha_p |pq|) \\
&= \operatorname{gcd}(\alpha_q |p|^2, \alpha_q |q|^2) = \alpha_q \operatorname{gcd}(|p|^2, |q|^2) \quad \text{and} \\
|\operatorname{grcd}(q, |pq|)|^2 &= \operatorname{gcd}(|q|^2, |pq|) = \operatorname{gcd}(\alpha_p |q|^2, \alpha_q |pq|) \\
&= \operatorname{gcd}(\alpha_p |q|^2, \alpha_p |p|^2) = \alpha_p \operatorname{gcd}(|q|^2, |p|^2).
\end{aligned}$$

□

PROPOSITION 3.25. *Let (p, q) be a primitive admissible pair of icosians and (p_α, q_α) its extension pair. Then,*

$$\Sigma_{\mathbb{I}}(R(p, q))\Sigma_{\mathbb{I}}(R(\overline{p}, q)) = \mathrm{N}(\mathrm{nr}(p_\alpha)^2) = \mathrm{N}(\mathrm{nr}(q_\alpha)^2).$$

PROOF. By Theorem 3.20 it is clear that

$$p_\alpha\mathbb{I} \subset p_\alpha\mathbb{I} + \mathbb{I}q_\alpha = \mathbb{I} \cap \tfrac{1}{|pq|}p\mathbb{I}q \subset \mathbb{I}.$$

and by Lemma 1.9 we know that $[\mathbb{I} : p_\alpha\mathbb{I}] = \mathrm{N}(\mathrm{nr}(p_\alpha)^2)$. Moreover,

$$[(p_\alpha\mathbb{I}+\mathbb{I}q_\alpha) : p_\alpha\mathbb{I}] = [(\mathbb{I}+p_\alpha^{-1}\mathbb{I}q_\alpha) : \mathbb{I}] = [(\mathbb{I}+\frac{\overline{p_\alpha}}{|p_\alpha q_\alpha|}\mathbb{I}q_\alpha) : \mathbb{I}] = [(\mathbb{I}+R(\overline{p}_\alpha, q_\alpha)\mathbb{I}) : \mathbb{I}]$$

and, the first isomorphism theorem, see for example [**53**, p. 17], implies that

$$\begin{aligned}[(\mathbb{I} + R(\overline{p}_\alpha, q_\alpha)\mathbb{I}) : \mathbb{I}] &= [R(\overline{p}_\alpha, q_\alpha)\mathbb{I} : (R(\overline{p}_\alpha, q_\alpha)\mathbb{I} \cap \mathbb{I})] \\ &= [\mathbb{I} : (\mathbb{I} \cap R(\overline{p}_\alpha, q_\alpha)\mathbb{I})] = \Sigma_{\mathbb{I}}(R(\overline{p}_\alpha, q_\alpha)).\end{aligned}$$

So we have

$$\mathrm{N}(\mathrm{nr}(p_\alpha)^2) = [\mathbb{I} : (p_\alpha\mathbb{I}+\mathbb{I}q_\alpha)][(p_\alpha\mathbb{I}+\mathbb{I}q_\alpha) : p_\alpha\mathbb{I}] = \Sigma_{\mathbb{I}}(R(p_\alpha, q_\alpha))\Sigma_{\mathbb{I}}(R(\overline{p}_\alpha, q_\alpha))$$

and $\mathrm{N}(\mathrm{nr}(p_\alpha)^2) = \mathrm{N}(\mathrm{nr}(q_\alpha)^2)$ see (3.17).

□

PROPOSITION 3.26. *Let (p, q) be a primitive admissible pair of icosians and (p_α, q_α) its extension. Define $g_p, g_q, h_p, h_q, k_p, k_q, r_p, r_q \in \mathbb{I}$, such that*

$$\begin{aligned} &g_p = \mathrm{glcd}(p, \alpha_q) = \mathrm{glcd}(p_\alpha, \alpha_q), \quad g_q = \mathrm{grcd}(q, \alpha_p) = \mathrm{grcd}(q_\alpha, \alpha_p), \\ &p = g_p h_p, \quad q = h_q g_q, \\ &k_p = \mathrm{glcd}(h_p, |h_q|^2), \quad k_q = \mathrm{grcd}(h_q, |h_p|^2) \\ &p = g_p k_p r_p \quad \textit{and} \quad q = r_q k_q g_q. \end{aligned} \tag{3.22}$$

Then,

$$\mathrm{gcd}(|h_p|^2, |h_q|^2) = \mathrm{gcd}(|p|^2, |q|^2) = |k_p|^2 = |k_q|^2,$$

and

$$g_p k_p = \mathrm{glcd}(p, |pq|), \quad k_q g_q = \mathrm{grcd}(q, |pq|).$$

Moreover, g_p, g_q, k_p, k_q, r_p and r_q are, up to units, uniquely identified as the factors

$$p = g_p k_p r_p \quad \textit{and} \quad q = r_q k_q g_q$$

such that $|g_p|^2 = |r_p|^2 = \alpha_q, |g_q|^2 = |r_q|^2 = \alpha_p$ and $|k_p|^2 = |k_q|^2 = \gcd(|p|^2, |q|^2)$.

Proof. Note that by Corollary 1.25 we have $|g_p|^2 = \alpha_q$ and $|g_q|^2 = \alpha_p$. Since α_p and α_q are relatively prime by definition this implies that $\gcd(|p|^2, |q|^2)| = \gcd(|h_p|^2, |h_q|^2) =: g$. Moreover, Corollary 1.25 reveals that $|\operatorname{glcd}(h_p, g)|^2 = g = |\operatorname{grcd}(h_q, g)|^2$ and $|k_p|^2 = |\operatorname{glcd}(h_p, |h_q|^2)|^2 = g = |\operatorname{grcd}(h_q, |h_p|^2)|^2 = |k_q|^2$. This proves the first claim.

Considering the definitions of the involved icosians, we see immediately that $g_p k_p$ divides p and $g_p |h_q|^2$. Since g_p divides α_q, it follows with (3.18) that $g_p k_p$ divides $\alpha_q |h_q|^2 = \frac{\alpha_q}{\alpha_p} |q|^2 = |pq|$. Thus $g_p k_p$ divides $\operatorname{glcd}(p, |pq|)$. By Lemma 3.24 we have

$$|\operatorname{glcd}(p, |pq|)|^2 = \alpha_q \gcd(|p|^2, |q|^2) = |g_p|^2 |k_p|^2,$$

which implies that $g_p k_p = \operatorname{glcd}(p, |pq|)$. Similarly, we see that $k_q g_q$ divides q and $|h_p|^2 g_q$. Since g_q divides α_p, it follows with (3.18) that $k_q g_q$ divides $\alpha_p |h_p|^2 = \frac{\alpha_p}{\alpha_q} |p|^2 = |pq|$. Thus $k_q g_q$ divides $\operatorname{grcd}(q, |pq|)$. By Lemma 3.24 we have

$$|\operatorname{grcd}(q, |pq|)|^2 = \alpha_p \gcd(|p|^2, |q|^2) = |k_q|^2 |g_q|^2,$$

which implies that $k_q g_q = \operatorname{grcd}(q, |pq|)$. Thus we have proved the second claim.

To prove the third claim, recall from (3.16) that $|p|^2 = \alpha_q^2 \gcd(|p|^2, |q|^2)$ and $|q|^2 = \alpha_p^2 \gcd(|p|^2, |q|^2)$. This implies that $|r_p|^2 = \alpha_q$ and $|r_q|^2 = \alpha_p$. Hence the icosians defined in (3.22) fulfil the third part of the claim. Moreover, as divisors of primitive icosians, with a fixed order of the elements in the prime factorisation of their norm, they are, up to units, uniquely identified by their norm. This completes the proof. □

PROPOSITION 3.27. *Let $p, q \in \mathbb{I}$ be primitive and let $\alpha \in \mathbb{Z}[\tau]$ such that $\alpha\mathbb{I}q \subset p\mathbb{I}$. Then,* $\mathrm{nr}(p)$ *divides* α*. Moreover, if $\beta \in \mathbb{Z}[\tau]$ denotes the element with minimal norm* $\mathrm{N}(\beta)$ *such that $\beta\mathbb{I}q \subset p\mathbb{I}$, then $\beta = \mathrm{nr}(p)$, up to units of* $\mathbb{Z}[\tau]$*.*

PROOF. If p and q are units of $\mathbb{I}$, the claim is clear. So assume that either p or q is not a unit of $\mathbb{I}$. $\mathbb{I}q \nsubseteq p\mathbb{I}$, since otherwise Lemma 1.20 implies that $\mathbb{I} = \mathbb{I}q\mathbb{I} \subset p\mathbb{I}$ which means that p and q are units of $\mathbb{I}$.

Since $\mathbb{Z}[\tau]$ is a Euclidean domain there are $\gamma, \varrho \in \mathbb{Z}[\tau]$ such that $\alpha = \gamma\beta + \varrho$ and $0 \leq \mathrm{N}(\varrho) < \mathrm{N}(\beta)$. Consequently, $(\alpha - \gamma\beta)\mathbb{I}q = \varrho\mathbb{I}q \subset p\mathbb{I}$, which implies that $\varrho = 0$ due to the definition of β. Hence β is a divisor of α. In particular, β divides $\mathrm{nr}(p)$, as obviously $\mathrm{nr}(p)\mathbb{I}q \subset p\mathbb{I}$.

Assume that β is a proper divisor of $\mathrm{nr}(p)$. This implies by Theorem 1.23 that $g := \mathrm{glcd}(\beta, p)$ is a proper left divisor of β and p. So there are $h_p, h_\beta \in \mathbb{I}$, which are not units in $\mathbb{I}$, such that $p = gh_p, \beta = gh_\beta$ and $h_p\mathbb{I} + h_\beta\mathbb{I} = \mathbb{I}$. Hence $h_\beta\mathbb{I}q \subset h_p\mathbb{I}$ and $\mathbb{I}q = (h_p\mathbb{I}q + h_\beta\mathbb{I}q) \subset h_p\mathbb{I}$, which implies that $\mathbb{I}q\mathbb{I} \subset h_p\mathbb{I}$. By Lemma 1.20 the two-sided ideal $\mathbb{I}q\mathbb{I} = \mathbb{I}$ and therefore $\mathbb{I} \subset h_p\mathbb{I}$, which contradicts the fact that h_p is not a unit in $\mathbb{I}$. Thus $\beta = \mathrm{nr}(p)$. □

COROLLARY 3.28. *If $p, q \in \mathbb{I}$ are primitive, then*

$$\mathrm{N}(\mathrm{nr}(p)) \text{ divides } [(p\mathbb{I} + \mathbb{I}q) : p\mathbb{I}],$$

i.e. there are at least $\mathrm{N}(\mathrm{nr}(p))$ *cosets of $p\mathbb{I}$ in $p\mathbb{I} + \mathbb{I}q$.*

PROOF. Let $\beta := [(p\mathbb{I} + \mathbb{I}q) : p\mathbb{I}]_K$, then by Lemma 1.6 we know that $\beta(p\mathbb{I} + \mathbb{I}q) \subset p\mathbb{I}$ and hence $\beta\mathbb{I}q \subset p\mathbb{I}$. Consequently, by Proposition 3.27, β is a multiple of $\mathrm{nr}(p)$ and hence we conclude with Lemma 1.7 that $\mathrm{N}(\beta)$ is a multiple of $\mathrm{N}(\mathrm{nr}(p))$. □

THEOREM 3.29. *Let (p, q) be a primitive admissible pair of icosians and (p_α, q_α) its extension. Then,*

$$\Sigma_{\mathbb{I}}(p, q) = \mathrm{N}(\mathrm{lcm}(|p|^2, |q|^2)) = \mathrm{N}(|p_\alpha|^2) = \mathrm{N}(|q_\alpha|^2) = \mathrm{N}(\alpha_p\alpha_q|pq|).$$

PROOF. Recall from Theorem 3.20 that $\Sigma_{\mathbb{I}}(p,q) = [\mathbb{I} : (p_\alpha\mathbb{I} + \mathbb{I}q_\alpha)]$. Considering (3.17) we only need to show that $\Sigma_{\mathbb{I}}(p,q) = \mathrm{N}(\mathrm{nr}(p_\alpha))$. Note that $(\overline{p},q)$ is a primitive admissible pair of icosians, too. If we show that

$$\Sigma_{\mathbb{I}}(p,q) \text{ divides } \mathrm{N}(\mathrm{nr}(p_\alpha)), \tag{3.23}$$

this will imply for the pair $(\overline{p},q)$ that $\Sigma_{\mathbb{I}}(\overline{p},q)$ divides $\mathrm{N}(\mathrm{nr}(\overline{p}_\alpha))$. So there will be $m,n \in \mathbb{N}$ such that $\Sigma_{\mathbb{I}}(p,q)m = \mathrm{N}(\mathrm{nr}(p_\alpha)) = \mathrm{N}(\mathrm{nr}(\overline{p_\alpha})) = \Sigma_{\mathbb{I}}(\overline{p},q)n$ and Proposition 3.25 will imply that $m = n = 1$. Hence, it is sufficient to show (3.23).

If p_α and q_α are units, the statement (3.23) is clear, due to the characterisation of units in $\mathbb{I}$, see (1.22). So assume that either p_α or q_α is not a unit. Since $|p_\alpha|^2 = |q_\alpha|^2$, see (3.17), this means that neither p_α nor q_α is a unit of $\mathbb{I}$.

First, suppose that $\alpha_p = \alpha_q = 1$. Hence p_α and q_α are primitive and Corollary 3.28 implies that $\mathrm{N}(\mathrm{nr}(p_\alpha))$ divides $[(p_\alpha\mathbb{I} + \mathbb{I}q_\alpha) : p_\alpha\mathbb{I}]$. Since by Lemma 1.9

$$[\mathbb{I} : (p_\alpha\mathbb{I} + \mathbb{I}q_\alpha)][(p_\alpha\mathbb{I} + \mathbb{I}q_\alpha) : p_\alpha\mathbb{I}] = \mathrm{N}(\mathrm{nr}(p_\alpha))^2,$$

this proves (3.23) and hence $[\mathbb{I} : (p_\alpha\mathbb{I} + \mathbb{I}q_\alpha)] = \mathrm{N}(\mathrm{nr}(p_\alpha))$.

Now, we use Proposition 3.26 and its notation for the reduction of the general case to the case $\alpha_p = \alpha_q = 1$. Note that

$$\begin{aligned} p_\alpha\mathbb{I} + \mathbb{I}q_\alpha &= p_\alpha\mathbb{I} + p_\alpha\mathbb{I}q + p\mathbb{I}q_\alpha + \mathbb{I}q_\alpha \\ &= p_\alpha\mathbb{I} + p\mathbb{I}q_\alpha + p_\alpha\mathbb{I}q + \mathbb{I}q_\alpha \\ &= p\mathbb{I}(\alpha_p\mathbb{I} + \mathbb{I}q_\alpha) + (\alpha_q\mathbb{I} + \mathbb{I}p_\alpha)\mathbb{I}q \\ &= p\mathbb{I}g_q + g_p\mathbb{I}q \;=\; g_p(h_p\mathbb{I} + \mathbb{I}h_q)g_q. \end{aligned}$$

Thus Lemma 1.9 implies that

$$[\mathbb{I} : (p_\alpha\mathbb{I} + \mathbb{I}q_\alpha)] = \mathrm{N}(\alpha_p^2\alpha_q^2)[\mathbb{I} : (h_p\mathbb{I} + \mathbb{I}h_q)].$$

Since $|h_p|^2 = \frac{|p|^2}{\alpha_q} \neq \frac{|q|^2}{\alpha_p} = |h_q|^2$, we cannot apply the special case yet. Therefore we need another representation of the subgroup $(h_p\mathbb{I} + \mathbb{I}h_q)$. Note that

$$\begin{aligned} h_p\mathbb{I} + \mathbb{I}h_q &= h_p\mathbb{I} + \mathbb{I}|h_p|^2 + |h_q|^2\mathbb{I} + \mathbb{I}h_q \\ &= (h_p\mathbb{I} + |h_q|^2\mathbb{I}) + (\mathbb{I}h_q + \mathbb{I}|h_p|^2) \\ &= k_p\mathbb{I} + \mathbb{I}k_q. \end{aligned} \tag{3.24}$$

where $|k_p|^2 = |k_q|^2 = \gcd(|p|^2, |q|^2) = \frac{|p|^2}{\alpha_q^2} = \frac{|q|^2}{\alpha_p^2}$. Finally, the application of the special case gives

$$[\mathbb{I} : (p_\alpha\mathbb{I} + \mathbb{I}q_\alpha)] = \mathrm{N}(\alpha_q^2\alpha_p^2)[\mathbb{I} : (k_p\mathbb{I} + \mathbb{I}k_q)] = \mathrm{N}(\alpha_q^2\alpha_p^2\tfrac{|p|^2}{\alpha_q^2}) = \mathrm{N}(|p_\alpha|^2).$$

□

The proof of Theorem 3.29 implies even more. Equation (3.24) shows that

$$p_\alpha\mathbb{I} + \mathbb{I}q_\alpha = g_p(k_p\mathbb{I} + \mathbb{I}k_q)g_q$$

depends only on g_p, g_q, k_p and k_q. With Proposition 3.26 this leads to following representation of a CSM of $\mathbb{I}$.

Corollary 3.30. *Let (p, q) be a primitive admissible pair of icosians. Decompose $p = g_pk_pr_p$ and $q = r_qk_qg_q$, such that $|g_p|^2 = |r_p|^2 = \alpha_q$, $|g_q|^2 = |r_q|^2 = \alpha_p$ and $|k_p|^2 = |k_q|^2 = \gcd(|p|^2, |q|^2)$. Then*

$$\mathbb{I} \cap \frac{p\mathbb{I}q}{|pq|} = g_p(k_p\mathbb{I} + \mathbb{I}k_q)g_q.$$

The following two corollaries give sufficient conditions for the equality of two CSMs of $\mathbb{I}$.

Corollary 3.31. *Let (p_1, q_1) and (p_2, q_2) be two primitive admissible pairs of icosians, such that $|p_1q_1| = |p_2q_2|$, $\alpha_{p_1} = \alpha_{p_2}$ and $\alpha_{q_1} = \alpha_{q_2}$. If $\mathrm{glcd}(p_1, |p_1q_1|) = \mathrm{glcd}(p_2, |p_2q_2|)$ and $\mathrm{grcd}(q_1, |q_1p_1|) = \mathrm{grcd}(q_2, |q_2p_2|)$, then*

$$\mathbb{I} \cap \tfrac{p_1\mathbb{I}q_1}{|p_1q_1|} = \mathbb{I} \cap \tfrac{p_2\mathbb{I}q_2}{|p_2q_2|}.$$

PROOF. For $i \in \{1,2\}$ decompose $p_i = g_{p_i} k_{p_i} r_{p_i}$ and $q_i = r_{p_i} k_{p_i} q_{p_i}$, such that $|g_{p_i}|^2 = |r_{p_i}|^2 = \alpha_{q_i}$, $|g_{q_i}|^2 = |r_{q_i}|^2 = \alpha_{p_i}$ and $|k_{p_i}|^2 = |k_{q_i}|^2 = \gcd(|p_i|^2, |q_i|^2)$. By Corollary 3.30, we only have to show that

$$g_{p_1}(k_{p_1}\mathbb{I} + \mathbb{I}k_{q_1})g_{q_1} = g_{p_2}(k_{p_2}\mathbb{I} + \mathbb{I}k_{q_2})g_{q_2}.$$

By the assumptions and Proposition 3.26

$$g_{p_1} k_{p_1} = \operatorname{glcd}(p_1, |p_1 q_1|) = \operatorname{glcd}(p_2, |p_2 q_2|) = g_{p_2} k_{p_2}.$$

Since $\alpha_{q_1} = \alpha_{q_2}$ divides $|p_1 q_1| = \alpha_{p_1}\alpha_{q_1} \gcd(|p_1|, |q_1|) = |p_2 q_2|$, this implies that $g_{p_1} = \operatorname{glcd}(p_1, \alpha_{q_1}) = \operatorname{glcd}(p_2, \alpha_{q_2}) = g_{p_2}$.

Similarly, again by assumptions and Proposition 3.26

$$k_{q_1} g_{q_1} = \operatorname{grcd}(q_1, |p_1 q_1|) = \operatorname{grcd}(q_2, |p_2 q_2|) = k_{q_2} g_{q_2}.$$

Since $\alpha_{p_1} = \alpha_{p_2}$ divides $|p_1 q_1| = \alpha_{p_1}\alpha_{q_1} \gcd(|p_1|, |q_1|) = |p_2 q_2|$, this implies that $g_{q_1} = \operatorname{grcd}(q_1, \alpha_{p_1}) = \operatorname{grcd}(q_2, \alpha_{p_2}) = g_{q_2}$. □

COROLLARY 3.32. *Let (p_1, q_1) and (p_2, q_2) be two primitive admissible pairs of icosians, such that $|p_1 q_1| = |p_2 q_2|$ and $\operatorname{lcm}(|p_1|^2, |q_1|^2) = \operatorname{lcm}(|p_2|^2, |q_2|^2)$. If $\operatorname{glcd}(p_1, |p_1 q_1|) = \operatorname{glcd}(p_2, |p_2 q_2|)$ and $\operatorname{grcd}(q_1, |q_1 p_1|) = \operatorname{grcd}(q_2, |q_2 p_2|)$, then*

$$\mathbb{I} \cap \tfrac{p_1 \mathbb{I} q_1}{|p_1 q_1|} = \mathbb{I} \cap \tfrac{p_2 \mathbb{I} q_2}{|p_2 q_2|}.$$

PROOF. Due to the assumptions and the fact that $\operatorname{lcm}(|p_i|^2, |q_i|^2) = \frac{|p_i q_i|^2}{\gcd(|p_i|^2, |q_i|^2)}$ for $i \in \{1,2\}$, it is clear that $\gcd(|p_1|^2, |q_1|^2) = \gcd(|p_2|^2, |q_2|^2)$. For $i = 1, 2$ decompose $p_i = g_{p_i} k_{p_i} r_{p_i}$ and $q_i = r_{p_i} k_{p_i} q_{p_i}$, such that $|g_{p_i}|^2 = |r_{p_i}|^2 = \alpha_{q_i}$, $|g_{q_i}|^2 = |r_{q_i}|^2 = \alpha_{p_i}$ and $|k_{p_i}|^2 = |k_{q_i}|^2 = \gcd(|p_i|^2, |q_i|^2)$. By Proposition 3.26 we know that

$$g_{p_1} k_{p_1} = \operatorname{glcd}(p_1, |p_1 q_1|) = \operatorname{glcd}(p_2, |p_2 q_2|) = g_{p_2} k_{p_2}.$$

which implies that

$$\alpha_{q_1} \gcd(|p_1|^2, |q_1|^2) = |g_{p_1}|^2 |k_{q_1}|^2 = |g_{p_2}|^2 |k_{p_2}|^2 = \alpha_{q_2} \gcd(|p_2|^2, |q_2|^2).$$

Hence $\alpha_{q_1} = \alpha_{q_2}$. Completely analogously it follows that $\alpha_{p_1} = \alpha_{p_2}$ and we can apply Corollary 3.31. □

A first step in proving the converse statement of Corollary 3.32 is

LEMMA 3.33. *Let (p_1, q_1) and (p_2, q_2) be two primitive admissible pairs of icosians. If*

$$\mathbb{I} \cap \tfrac{p_1 \mathbb{I} q_1}{|p_1 q_1|} = \mathbb{I} \cap \tfrac{p_2 \mathbb{I} q_2}{|p_2 q_2|},$$

then $|p_1 q_1| = |p_2 q_2|$ and $\mathrm{N}(\mathrm{lcm}(|p_1|^2, |q_1|^2) = \mathrm{N}(\mathrm{lcm}(|p_2|^2, |q_2|^2)$, i.e. denominator and coincidence index are the same.

PROOF. Clearly, $|p_1 q_1| \mathbb{I} \subset \mathbb{I}$ and $|p_1 q_1|^2 \mathbb{I} \subset p_1 \mathbb{I} q_1$. Hence $|p_1 q_1| \mathbb{I} \subset \mathbb{I} \cap \frac{p_1 \mathbb{I} q_1}{|p_1 q_1|} = \mathbb{I} \cap \frac{p_2 \mathbb{I} q_2}{|p_2 q_2|}$. This implies directly that $\frac{|p_1 q_1|}{|p_2 q_2|} \overline{p_2} \mathbb{I} \overline{q_2} \subset \frac{\overline{p_2} \mathbb{I} \overline{q_2}}{|p_2 q_2|} \cap \mathbb{I}$, i.e. $|p_1 q_1|$ is a multiple of $\mathrm{den}_{\mathbb{I}}(R(\overline{p_2}, \overline{q_2}))$. As $(\overline{p_2}, \overline{q_2})$ is also a primitive admissible pair of icosians, this means that $\mathrm{den}_{\mathbb{I}}(R(\overline{p_2}, \overline{q_2})) = |\overline{p_2} \overline{q_2}| = |p_2 q_2|$. Exchanging the roles of $|p_2 q_2|$ and $|p_1 q_1|$ gives the claim. By Theorem 3.29 we know that $\Sigma_{\mathbb{I}}(p_1, q_1) = \mathrm{N}(\mathrm{lcm}(|p_1|^2, |q_1|^2)) = \mathrm{N}(\mathrm{lcm}(|p_2|^2, |q_2|^2)) = \Sigma_{\mathbb{I}}(p_1, q_1)$. □

For later use we gather some additional information on the index of $[\mathbb{I} : (r\mathbb{I} + \mathbb{I}s)]$ for arbitrary $r, s \in \mathbb{I}$. We first consider the case where r, s are both primitive.

LEMMA 3.34. *If $r, s \in \mathbb{I}$ are primitive, then*

$$[\mathbb{I} : (r\mathbb{I} + \mathbb{I}s)]_K \quad \textit{divides} \quad \gcd(|r|^2, |s|^2).$$

PROOF. If r and s are units of $\mathbb{I}$, the claim is clear. So assume that either r or s is not a unit of $\mathbb{I}$. Since $r\mathbb{I} \subset (r\mathbb{I} + \mathbb{I}s)$, we know by Lemmas 1.8 that,

$$[\mathbb{I} : (r\mathbb{I} + \mathbb{I}s)]_K \, [(r\mathbb{I} + \mathbb{I}s) : r\mathbb{I}]_K = \mathrm{nr}(r)^2.$$

Moreover, for $\beta = [(r\mathbb{I} + \mathbb{I}s) : r\mathbb{I}]_K$ Lemma 1.6 implies that $\beta(r\mathbb{I} + \mathbb{I}s) \subset r\mathbb{I}$ and thus $\beta\mathbb{I}s \subset r\mathbb{I}$. Consequently, by Proposition 3.27 $\mathrm{nr}(r)$ divides β and hence $[\mathbb{I} : (r\mathbb{I} + \mathbb{I}s)]_K$ divides $\mathrm{nr}(r)$. Similarly, $[\mathbb{I} : (r\mathbb{I} + \mathbb{I}s)]_K$ divides $\mathrm{nr}(s)$ and thus $[\mathbb{I} : (r\mathbb{I} + \mathbb{I}s)]_K$ divides $\gcd(\mathrm{nr}(r), \mathrm{nr}(s))$. □

LEMMA 3.35. *If $r, s \in \mathbb{I}$ are primitive and $\beta, \gamma \in \mathbb{Z}[\tau]$ are relatively prime, then*

$$[\mathbb{I} : (\beta r\mathbb{I} + \mathbb{I}\gamma s)] \text{ divides } \mathrm{N}(\gamma_r^2\beta_s^2 \gcd(\tfrac{|r|^2}{\gamma_r}, \tfrac{|s|^2}{\beta_s})).$$

where $\gamma_r := |\operatorname{glcd}(r, \gamma)|^2$ and $\beta_s := |\operatorname{grcd}(s, \beta)|^2$.

PROOF. Define $g_r := \operatorname{glcd}(r, \gamma)$ and $g_s := \operatorname{grcd}(s, \beta)$ and note that

$$\beta r\mathbb{I}s + \gamma r\mathbb{I}s = r\mathbb{I}s \subset \beta r\mathbb{I} + \mathbb{I}\gamma s.$$

The decomposition $r = g_r k_r, s = k_s g_s$ leads to

$$\begin{aligned}\beta r\mathbb{I} + \mathbb{I}\gamma s &= r\mathbb{I}\beta + r\mathbb{I}s + \gamma\mathbb{I}s + r\mathbb{I}s = r\mathbb{I}(\beta + s) + (\gamma + r)\mathbb{I}s \\ &= r\mathbb{I}g_s + g_r\mathbb{I}s = g_r(k_r\mathbb{I} + \mathbb{I}k_s)g_s\end{aligned}$$

With Lemma 1.9 we infer that

$$[\mathbb{I} : (\beta r\mathbb{I}+\mathbb{I}\gamma s)] = \mathrm{N}(\mathrm{nr}(g_r)^2)\,\mathrm{N}(\mathrm{nr}(g_s)^2)[\mathbb{I} : (k_r\mathbb{I}+\mathbb{I}k_s)] = \mathrm{N}(\gamma_r^2\beta_s^2)[\mathbb{I} : (k_r\mathbb{I}+\mathbb{I}k_s)].$$

Since k_r and k_s are primitive as factors of primitive icosians, Lemma 3.34 tells us that $[\mathbb{I} : (\beta r\mathbb{I} + \mathbb{I}\gamma s)]$ divides $\mathrm{N}(\gamma_r^2\beta_s^2)\,\mathrm{N}(\gcd(|k_r|^2, |k_s|^2))$. Observing that $|k_r|^2 = \frac{|r|^2}{\gamma_r}$ and $|k_s|^2 = \frac{|s|^2}{\beta_s}$ completes the proof. □

If we apply the previous lemma to a primitive admissible pair of icosians (p, q), where (p_α, q_α) denotes its extension pair, we get that $[\mathbb{I} : (p_\alpha\mathbb{I} + \mathbb{I}q_\alpha)]$ divides

$$\mathrm{N}(\alpha_p^2\alpha_q^2 \gcd(\tfrac{|p|^2}{\alpha_q}, \tfrac{|q|^2}{\alpha_p})) = \mathrm{N}(\alpha_p^2\alpha_q^2 \gcd(|p|^2, |q|^2)) = \mathrm{N}(\operatorname{lcm}(|p|^2, |q|^2)),$$

which is consistent with Theorem 3.29.

Now, we have everything we need to prove the converse statement of Corollary 3.32, and thus we get a necessary and sufficient condition for the equality of two CSM of $\mathbb{I}$ that are parameterised by different pairs of icosians.

THEOREM 3.36. *Let* (p_1, q_1) *and* (p_2, q_2) *be two primitive admissible pairs of icosians. Then,*

$$\mathbb{I} \cap \tfrac{q_1\mathbb{I}p_1}{|p_1q_1|} = \mathbb{I} \cap \tfrac{q_2\mathbb{I}p_2}{|p_2q_2|}$$

if and only if

$$|p_1q_1| = |p_2q_2|,\ \mathrm{lcm}(|p_1|^2, |q_1|^2) = \mathrm{lcm}(|p_2|^2, |q_2|^2),\ \mathrm{glcd}(p_1, |p_1q_1|) = \mathrm{glcd}(p_2, |p_2q_2|) \text{ and } \mathrm{grcd}(q_1, |q_1p_1|) = \mathrm{grcd}(q_2, |q_2p_2|).$$

PROOF. With Corollary 3.32 and Lemma 3.33 it only remains to show that $\mathbb{I} \cap \frac{q_1\mathbb{I}p_1}{|p_1q_1|} = \mathbb{I} \cap \frac{q_2\mathbb{I}p_2}{|p_2q_2|}$, $|p_1q_1| = |p_2q_2|$ and $\mathrm{N}(\mathrm{lcm}(|p_1|^2, |q_1|^2)) = \mathrm{N}(\mathrm{lcm}(|p_2|^2, |q_2|^2))$, implies

$$\mathrm{lcm}(|p_1|^2, |q_1|^2) = \mathrm{lcm}(|p_2|^2, |q_2|^2), \mathrm{glcd}(p_1, |p_1q_1|) = \mathrm{glcd}(p_2, |p_2q_2|) \text{ and } \mathrm{grcd}(q_1, |q_1p_1|) = \mathrm{glcd}(q_2, |q_2p_2|).$$

We infer from Corollary 3.22 that

$$p_{1\ell}\mathbb{I} + \mathbb{I}q_{1r} = \mathbb{I} \cap \tfrac{q_1\mathbb{I}p_1}{|p_1q_1|} = \mathbb{I} \cap \tfrac{q_2\mathbb{I}p_2}{|p_2q_2|} = p_{2\ell}\mathbb{I} + \mathbb{I}q_{2r},$$

where $p_{i\ell} = \mathrm{glcd}(p_{i\alpha}, |p_1q_1|) = \alpha_{p_i}\,\mathrm{glcd}(p_i, \frac{|p_1q_1|}{\alpha_{p_i}})$ and $q_{ir} = \mathrm{grcd}(q_{i\alpha}, |p_1q_1|) = \alpha_{q_i}\,\mathrm{grcd}(p_i, \frac{|p_1q_1|}{\alpha_{q_i}})$ for $i \in \{1, 2\}$. Clearly,

$$p_{1\ell}\mathbb{I} + \mathbb{I}q_{1r} = p_{1\ell}\mathbb{I} + \mathbb{I}q_{1r} + p_{2\ell}\mathbb{I} + \mathbb{I}q_{2r} = (p_{1\ell} + p_{2\ell})\mathbb{I} + \mathbb{I}(q_{1r} + q_{2r}) = \beta r\mathbb{I} + \mathbb{I}\gamma s,$$

where $\beta r := \mathrm{glcd}(p_{1\ell}, p_{2\ell})$ and $\gamma s := \mathrm{grcd}(q_{1r}, q_{2r})$ such that $r, s \in \mathbb{I}$ are primitive and $\beta, \gamma \in \mathbb{Z}[\tau]$. Note that $\beta = \gcd(\alpha_{p_1}, \alpha_{p_2})$, since p_1 and p_2 are primitive and therefore do not have any divisors in $\mathbb{Z}[\tau]$. Similarly, $\gamma = \gcd(\alpha_{q_1}, \alpha_{q_2})$. Since α_{p_1} and α_{q_1} are relatively prime, β and γ are

relatively prime, too. So we have

$$r = \operatorname{glcd}(\operatorname{glcd}(p_1, \tfrac{|p_1q_1|}{\alpha_{p_1}}), \operatorname{glcd}(p_2, \tfrac{|p_1q_1|}{\alpha_{p_2}})) \quad \text{and}$$
$$s = \operatorname{grcd}(\operatorname{grcd}(q_1, \tfrac{|p_1q_1|}{\alpha_{q_1}}), \operatorname{grcd}(q_2, \tfrac{|p_1q_1|}{\alpha_{q_2}})).$$

In this representation it is clear that r divides $\operatorname{glcd}(p_1, \frac{|p_1q_1|}{\alpha_{p_1}})$ and s divides $\operatorname{grcd}(q_1, \frac{|p_1q_1|}{\alpha_{q_1}})$. By (3.18) we know that $\frac{|p_1q_1|}{\alpha_{p_1}}$ divides $|p_1|^2$ and $\frac{|p_1q_1|}{\alpha_{q_1}}$ divides $|q_1|^2$. Hence we infer with Corollary 1.25 that $|r|^2$ divides $|\operatorname{glcd}(p_1, \frac{|p_1q_1|}{\alpha_{p_1}})|^2 = \frac{|p_1q_1|}{\alpha_{p_1}}$ and $|s|^2$ divides $|\operatorname{grcd}(q_1, \frac{|p_1q_1|}{\alpha_{q_1}})|^2 = \frac{|p_1q_1|}{\alpha_{q_1}}$. This means that

$$\beta|r|^2 \quad \text{and} \quad \gamma|s|^2 \text{ divide } |p_1q_1|. \tag{3.25}$$

Thus $\gcd(|r|^2, |s|^2)$ divides $\gcd(\frac{|p_1q_1|}{\beta}, \frac{|p_1q_1|}{\gamma}) = \frac{|p_1q_1|}{\beta\gamma}$ and

$$\beta\gamma \gcd(|r|^2, |s|^2) \text{ divides } |p_1q_1|. \tag{3.26}$$

An application of Lemma 3.35 reveals that

$$[\mathbb{I} : (\beta r\mathbb{I} + \mathbb{I}\gamma s)] \text{ divides } \mathrm{N}(\gamma_r^2\beta_s^2 \gcd(\tfrac{|r|^2}{\gamma_r}, \tfrac{|s|^2}{\beta_s})),$$

where $\gamma_r = |\operatorname{glcd}(r, \gamma)|^2 = \gcd(|r|^2, \gamma)$ and $\beta_s = |\operatorname{grcd}(s, \beta)|^2 = \gcd(|s|^2, \beta)$ by Corollary 1.25. Since γ_r divides γ, β_s divides β and $\gcd(\frac{|r|^2}{\gamma_r}, \frac{|s|^2}{\beta_s})$ divides $\gcd(|r|^2, |s|^2)$ it follows with (3.26) that $\beta_s^2\gamma_r^2 \gcd(\frac{|r|^2}{\gamma_r}, \frac{|s|^2}{\beta_s})$ divides $\beta_s\gamma_r|p_1q_1|$. Therefore we know that $[\mathbb{I} : (\beta r\mathbb{I} + \mathbb{I}\gamma s)]$ divides $\mathrm{N}(\beta_s\gamma_r|p_1q_1|)$.

By Theorem 3.29 we infer that $[\mathbb{I} : (\beta r\mathbb{I} + \mathbb{I}\gamma s)] = [\mathbb{I} : (p_{1\ell}\mathbb{I} + \mathbb{I}q_{1r})] = \mathrm{N}(\alpha_{p_1}\alpha_{q_1}|p_1q_1|)$, which implies that $\mathrm{N}(\alpha_{p_1}\alpha_{q_1})$ divides $\mathrm{N}(\beta_s\gamma_r)$. Since $\gcd(\alpha_{p_1}, \alpha_{q_1}) = 1 = \gcd(\beta_s, \gamma_r)$, we conclude that $\mathrm{N}(\alpha_{p_1})$ divides $\mathrm{N}(\beta_s)$ and $\mathrm{N}(\alpha_{q_1})$ divides $\mathrm{N}(\gamma_r)$. If we take into account that β_s divides α_{p_1} and γ_r divides α_{q_1}, we see that necessarily

$$\beta_s = \beta = \alpha_{p_1} = \alpha_{p_2} \quad \text{and} \quad \gamma_r = \gamma = \alpha_{q_1} = \alpha_{q_2}.$$

This implies directly that

$$\operatorname{lcm}(|p_1|^2, |q_1|^2) = \alpha_{p_1}\alpha_{q_1}|p_1q_1| = \alpha_{p_2}\alpha_{q_2}|p_2q_2| = \operatorname{lcm}(|p_2|^2, |q_2|^2).$$

Another application of Lemma 3.35 tells us that

$$[\mathbb{I} : (\beta r\mathbb{I} + \mathbb{I}\gamma s)] = \mathrm{N}(\alpha_{p_1}\alpha_{q_1}|p_1q_1|) \text{ divides } \mathrm{N}(\alpha_{p_1}^2\alpha_{q_1}^2 \gcd(\tfrac{|r|^2}{\alpha_{q_1}}, \tfrac{|s|^2}{\alpha_{p_1}})),$$

which implies that $\mathrm{N}(\frac{|p_1q_1|}{\alpha_{p_1}\alpha_{q_1}})$ divides $\mathrm{N}(\gcd(\frac{|r|^2}{\alpha_{q_1}}, \frac{|s|^2}{\alpha_{p_1}}))$ and in particular

$$\mathrm{N}\left(\tfrac{|p_1q_1|}{\alpha_{p_1}}\right) \text{ divides } \mathrm{N}\left(|r|^2\right) \quad \text{and} \quad \mathrm{N}\left(\tfrac{|p_1q_1|}{\alpha_{q_1}}\right) \text{ divides } \mathrm{N}\left(|s|^2\right).$$

With (3.25) and Lemma 3.24 this shows that

$$|r|^2 = \tfrac{|p_1q_1|}{\alpha_{p_1}} = |\operatorname{glcd}(p_1, |p_1q_1|)|^2 \quad \text{and} \quad |s|^2 = \tfrac{|p_1q_1|}{\alpha_{q_1}} = |\operatorname{grcd}(q_1, |p_1q_1|)|^2.$$

Hence $r = \operatorname{glcd}(p_1, |p_1q_1|) = \operatorname{glcd}(p_2, |p_2q_2|)$ and $s = \operatorname{grcd}(q_1, |p_1q_1|)$ $= \operatorname{grcd}(q_2, |p_2q_2|)$, which completes the proof. □

3.4. Counting Coincidence Site Lattices of A_4

In this section we continue our analysis of the CSLs of the lattice L. Recall from Theorem 3.11 that the CSLs of L are precisely the lattices of the form $L \cap R(p)L$, where p is a primitive and admissible icosian. The results of the previous section lead to the following formula for its coincidence index.

THEOREM 3.37. *Let $p \in \mathbb{I}$ be primitive and admissible, then*

$$\Sigma_L(p) \;=\; \mathrm{nr}(p_\alpha) \;=\; \operatorname{lcm}(\mathrm{nr}(p), \mathrm{nr}(p)')\,.$$

PROOF. By Lemma 3.18 we know that $\Sigma_L^2(p) = \Sigma_{L[\tau]}(p)$. Moreover, Theorem 3.23 tells us that $\Sigma_{L[\tau]}(p) = \Sigma_{\mathbb{I}}(p)$. Theorem 3.29 implies that $\Sigma_{\mathbb{I}}(p) = \Sigma_{\mathbb{I}}(R(p, \widetilde{p})) = \mathrm{N}(\operatorname{lcm}(\mathrm{nr}(p), \mathrm{nr}(\widetilde{p})) = \mathrm{N}(\operatorname{lcm}(\mathrm{nr}(p), \mathrm{nr}(p)'))$ and by (3.10), we know that $\mathrm{nr}(p_\alpha) = \operatorname{lcm}(\mathrm{nr}(p), \mathrm{nr}(p)')$ is always an element of $\mathbb{N}$. Hence $\Sigma_{\mathbb{I}}(p) = \mathrm{nr}(p_\alpha)^2$ and we conclude that $\Sigma_L(p) \;=\; \mathrm{nr}(p_\alpha)$. □

In analogy to Theorem 3.36 we want to find necessary and sufficient conditions for the equality of two CSLs of L that result from different rotations. The following lemma gives a first necessary condition.

LEMMA 3.38. *If $R_1, R_2 \in \mathrm{SOC}(L)$ such that $L \cap R_1 L = L \cap R_2 L$, then $\Sigma(R_1) = \Sigma(R_2)$ and $\mathrm{den}(R_1) = \mathrm{den}(R_2)$.*

PROOF. Lemma 3.9 implies that

$$\Sigma(R_1) = \Sigma(R_2) \quad \text{and} \quad \mathrm{den}(R_1^{-1}) = \mathrm{den}(R_2^{-1}).$$

For $i \in \{1, 2\}$ let $p_i \in \mathbb{I}$ be primitive, admissible and such that $R_i = R_i(p_i)$. As $R_i^{-1} = R(\overline{p}_i)$, it is clear that $\mathrm{den}(R_1) = |p_1 \widetilde{p_1}| = \mathrm{den}(R_1^{-1}) = \mathrm{den}(R_2^{-1}) = |p_2 \widetilde{p_2}| = \mathrm{den}(R_2)$. □

Recall from Proposition 1.15 the $\mathbb{Q}$-linear map $\varphi_+ \colon \mathbb{H}(K) \longrightarrow \mathbb{H}(K)$, defined by $\varphi_+(x) = x + \widetilde{x}$, and note that not only $\varphi_+(\mathbb{I}) = L$ but that

$$\varphi_+(p_\alpha \mathbb{I}) = L(p_\alpha) = L \cap \tfrac{1}{|\widetilde{p}p|} p L \widetilde{p} = (p_\alpha \mathbb{I} + \mathbb{I}\widetilde{p}_\alpha) \cap L, \tag{3.27}$$

see Theorem 3.16 and Corollary 3.17. So two CSLs of L are certainly equal if the corresponding CSMs of $\mathbb{I}$ are equal. Thus the following lemma is just an application of Theorem 3.36. We only need to observe with Lemma 1.27 that $\mathrm{glcd}(p_1, |p_1 \widetilde{p}_1|) = \mathrm{glcd}(p_2, |p_2 \widetilde{p}_2|)$ implies

$$\mathrm{grcd}(\widetilde{p}_1, |p_1 \widetilde{p}_1|) = (\mathrm{glcd}(p_1, |p_1 \widetilde{p}_1|))\widetilde{\ } = (\mathrm{glcd}(p_2, |p_2 \widetilde{p}_2|))\widetilde{\ } = \mathrm{grcd}(\widetilde{p}_2, |p_2 \widetilde{p}_2|),$$

and that $|p_1|^2 = |p_2|^2$ implies $\mathrm{lcm}(|p_1|^2, |\widetilde{p}_1|^2) = \mathrm{lcm}(|p_2|^2, |\widetilde{p}_2|^2)$, as well as, $|p_1 \widetilde{p}_1| = |p_2 \widetilde{p}_2|$ with Lemma 3.24.

LEMMA 3.39. *Let $p_1, p_2 \in \mathbb{I}$ be primitive and admissible, such that*

$$|p_1|^2 = |p_2|^2 \quad \textit{and} \quad \mathrm{glcd}(p_1, |p_1 \widetilde{p}_1|) = \mathrm{glcd}(p_2, |p_2 \widetilde{p}_2|).$$

Then, one has $L \cap \frac{1}{|p_1 \widetilde{p_1}|} p_1 L \widetilde{p_1} = L \cap \frac{1}{|p_2 \widetilde{p_2}|} p_2 L \widetilde{p_2}$.

It turns out that the converse is not true. However, by considering certain cases separately, we can derive necessary and sufficient conditions on the parameterising icosians so that they generate the same CSL of L.

THEOREM 3.40. *Let $p_1, p_2 \in \mathbb{I}$ be primitive, admissible and such that $|p_1|^2$ or $|p_2|^2$ is not divisible by 5. Then, one has*

$$L \cap \tfrac{1}{|p_1\widetilde{p_1}|} p_1 L\widetilde{p_1} = L \cap \tfrac{1}{|p_2\widetilde{p_2}|} p_2 L\widetilde{p_2}$$

if and only if

$$|p_1|^2 = |p_2|^2 \quad \textit{and} \quad \operatorname{glcd}(p_1, |p_1\widetilde{p_1}|) = \operatorname{glcd}(p_2, |p_2\widetilde{p_2}|).$$

PROOF. Considering Lemma 3.39 and (3.27), it only remains to show that

$$L(p_{1\alpha}) = \varphi_+(p_{1\alpha}\mathbb{I}) = \varphi_+(p_{2\alpha}\mathbb{I}) = L(p_{2\alpha})$$

implies $|p_1|^2 = |p_2|^2$ and $\operatorname{glcd}(p_1, |p_1\widetilde{p_1}|) = \operatorname{glcd}(p_2, |p_2\widetilde{p_2}|)$. If we deduce that

$$p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha} = p_{2\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{2\alpha},$$

this is delivered by an application of Theorem 3.36, since $|p_1\widetilde{p_1}| = |p_2\widetilde{p_2}|$ and $\operatorname{glcd}(p_1, |p_1\widetilde{p_1}|) = \operatorname{glcd}(p_2, |p_2\widetilde{p_2}|)$, implies with Lemma 3.24 that $\frac{1}{\alpha_{p_1}}|p_1\widetilde{p_1}| = \frac{1}{\alpha_{p_2}}|p_2\widetilde{p_2}|$ and hence $\alpha_{p_1}^2 = \alpha_{p_2}^2 = \frac{\operatorname{lcm}(|p_1|^2, |\widetilde{p_1}|^2)}{|p_1|^2} = \frac{\operatorname{lcm}(|p_2|^2, |\widetilde{p_2}|^2)}{|p_2|^2}$, which reveals with $\operatorname{lcm}(|p_1|^2, |\widetilde{p}_1|^2) = \operatorname{lcm}(|p_2|^2, |\widetilde{p}_2|^2)$ that $|p_1|^2 = |p_2|^2$.

With Lemma 3.18 and (3.27) it is clear that

$$\begin{aligned} &L(p_{1\alpha}) + \tau L(p_{1\alpha}) = \varphi_+(p_{1\alpha}\mathbb{I}) + \tau\varphi_+(p_{1\alpha}\mathbb{I}) \\ &= \varphi_+(p_{1\alpha}\mathbb{I}) \cap \varphi_+(p_{2\alpha}\mathbb{I}) + \tau(\varphi_+(p_{1\alpha}\mathbb{I}) \cap \varphi_+(p_{2\alpha}\mathbb{I})) \\ &\subset (p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha}) \cap (p_{2\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{2\alpha}) \\ &\subset p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha} \subset \mathbb{I} \end{aligned}$$

Considering $[\mathbb{I} : L[\tau]] = 5$, compare (1.28), and Theorem 3.23 we see that

(3.28)
$$[\mathbb{I} : L(p_{1\alpha}) + \tau L(p_{1\alpha})] = 5\Sigma_{L[\tau]}(p_1) = 5\Sigma_{L[\tau]}(p_2) = 5\Sigma_{\mathbb{I}}(p_1) = 5\Sigma_{\mathbb{I}}(p_2)$$

and hence $\Sigma_{\mathbb{I}}(p_1) = \Sigma_{\mathbb{I}}(p_2) = \mathrm{N}(\operatorname{lcm}(|p_1|^2, |\widetilde{p_1}|^2) = \mathrm{N}(\operatorname{lcm}(|p_2|^2, |\widetilde{p_2}|^2)$ by Theorem 3.29. Without loss of generality let $|p_1|^2$ be not divisible by 5. By

Lemma 3.12 this implies that $|p_1|^2$ and $|\widetilde{p_1}|^2$ are not divisible by $\sqrt{5}$ and hence $\mathrm{N}(\mathrm{lcm}(|p_1|^2, |\widetilde{p_1}|^2) = \Sigma_{\mathbb{I}}(p_1) = \Sigma_{\mathbb{I}}(p_2)$ is not divisible by 5.

Note that $[\mathbb{I} : (p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha}) \cap (p_{2\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{2\alpha})]$ must be a multiple of $\Sigma_{\mathbb{I}}(p_1)$ and a divisor of $5\Sigma_{\mathbb{I}}(p_1)$, i.e. it is either $\Sigma_{\mathbb{I}}(p_1)$ or $5\Sigma_{\mathbb{I}}(p_1)$. Assume that

$$[\mathbb{I} : (p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha}) \cap (p_{2\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{2\alpha})] = 5\Sigma_{\mathbb{I}}(p_1).$$

This implies with the first isomorphism theorem, see for example [**53**] that

$$\begin{aligned} &[(p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha}) : (p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha}) \cap (p_{2\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{2\alpha})] \\ &= [(p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha}) + (p_{2\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{2\alpha}) : (p_{2\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{2\alpha})] = 5. \end{aligned} \tag{3.29}$$

Moreover,

$$[\mathbb{I} : (p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha}) + (p_{2\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{2\alpha})] = \tfrac{\Sigma_{\mathbb{I}}(p_2)}{5}$$

which contradicts the fact that $\Sigma_{\mathbb{I}}(p_2) = \Sigma_{\mathbb{I}}(p_1)$ is not divisible by 5.

So we have necessarily

$$[\mathbb{I} : (p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha}) \cap (p_{2\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{2\alpha})] = \Sigma_{\mathbb{I}}(p_1)$$

and therefore

$$(p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha}) = (p_{1\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{1\alpha}) \cap (p_{2\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{2\alpha}) = (p_{2\alpha}\mathbb{I} + \mathbb{I}\widetilde{p}_{2\alpha}).$$

□

LEMMA 3.41. *Let $p_1, p_2 \in \mathbb{I}$ be primitive, admissible and such that $|p_1|^2$ or $|p_2|^2$ is a power of 5. Then,*

$$L \cap \tfrac{1}{|p_1\widetilde{p_1}|} p_1 L\widetilde{p_1} = L \cap \tfrac{1}{|p_2\widetilde{p_2}|} p_2 L\widetilde{p_2}$$

if and only if

$$|p_1|^2 = |p_2|^2 \quad \textit{and} \quad \mathrm{gcd}(p_1, \tfrac{|p_1\widetilde{p}_1|}{\sqrt{5}}) = \mathrm{gcd}(p_2, \tfrac{|p_2\widetilde{p}_2|}{\sqrt{5}}).$$

PROOF. Let $L \cap \frac{1}{|p_1\widetilde{p_1}|}p_1 L\widetilde{p_1} = L \cap \frac{1}{|p_2\widetilde{p_2}|}p_2 L\widetilde{p_2}$. If $|p_1|^2 = 5^m$ for some $m \in \mathbb{N}$, then $|\widetilde{p}_1|^2 = 5^m$. With Lemma 3.38 and Theorem 3.37 we see that

$$\Sigma_L(p_1) = \mathrm{lcm}(|p_1|^2, |\widetilde{p}_1|^2) = |p_1|^2 = 5^m = \Sigma_L(p_2) = \mathrm{lcm}(|p_2|^2, |\widetilde{p}_2|^2) = |p_2|^2.$$

Hence $\alpha_{p_1} = 1 = \alpha_{p_2}$, i.e. $p_1 = p_{1\alpha}, p_2 = p_{2\alpha}$ and $|p_1\widetilde{p_1}| = |p_1|^2, |p_2\widetilde{p_2}| = |p_2|^2$. We proceed now as in the proof of Theorem 3.40 and find that

$$[\mathbb{I} : (p_1\mathbb{I} + \mathbb{I}\widetilde{p}_1) \cap (p_2\mathbb{I} + \mathbb{I}\widetilde{p}_2)] = \Sigma_{\mathbb{I}}(p_1) \quad \text{or} \quad = 5\Sigma_{\mathbb{I}}(p_1). \tag{3.30}$$

In the first case, we follow the argumentation in the proof of Theorem 3.40 and conclude that $\mathrm{glcd}(p_1, |p_1\widetilde{p}_1|) = \mathrm{glcd}(p_2, |p_2\widetilde{p}_2|)$ and hence

$$\mathrm{glcd}(p_1, \tfrac{|p_1\widetilde{p}_1|}{\sqrt{5}}) = \mathrm{glcd}(p_2, \tfrac{|p_2\widetilde{p}_2|}{\sqrt{5}}).$$

In the second case we find, see (3.29), that

$$[(p_1\mathbb{I} + \mathbb{I}\widetilde{p}_1) + (p_2\mathbb{I} + \mathbb{I}\widetilde{p}_2) : (p_2\mathbb{I} + \mathbb{I}\widetilde{p}_2)] = 5$$

If $g = \mathrm{glcd}(p_1, p_2)$ we have $g\mathbb{I} + \mathbb{I}\widetilde{g} = (p_1\mathbb{I} + \mathbb{I}\widetilde{p}_1) + (p_2\mathbb{I} + \mathbb{I}\widetilde{p}_2)$ and since $g \in \mathbb{I}$ and $|g|^2 = (\sqrt{5})^\ell$, for some $\ell \in \mathbb{N}$, we have $|g\widetilde{g}| = (\sqrt{5})^{\frac{\ell}{2}}(\sqrt{5})^{\frac{\ell}{2}} = (\sqrt{5})^\ell \in \mathbb{Z}[\tau]$. Hence $(g, \widetilde{g})$ is $\mathbb{Z}[\tau]$-admissible. Note that g and $\widetilde{g}$ are, as factors of primitive icosians, primitive themselves. Moreover, $\alpha_g = \alpha_{\widetilde{g}} = 1$ so that according to Theorem 3.29 $[\mathbb{I} : g\mathbb{I} + \mathbb{I}\widetilde{g}] = \mathrm{N}(|g|^2)$, i.e. $\mathrm{N}(|p_2|^2) = \mathrm{N}(|g|^2)5$. Since $|g|^2\sqrt{5}$ divides $|p_2|^2$, this implies that up to units of $\mathbb{Z}[\tau]$ we have $|p_2|^2 = \sqrt{5}|g|^2$ and hence $\mathrm{glcd}(p_2, |g|^2) = \mathrm{glcd}(p_2, \frac{|p_2|^2}{\sqrt{5}})$. Symmetrically, we infer that $\mathrm{glcd}(p_1, |g|^2) = \mathrm{glcd}(p_1, \frac{|p_1|^2}{\sqrt{5}})$.

Conversely, let $|p_1|^2 = |p_2|^2 = (\sqrt{5})^r$, for $r \in 2\mathbb{N}$, and $\mathrm{glcd}(p_1, \frac{|p_1\widetilde{p}_1|}{\sqrt{5}}) = \mathrm{glcd}(p_2, \frac{|p_2\widetilde{p}_2|}{\sqrt{5}})$. Then either $p_1 = p_2$ and the proof is complete, or $p_1 \neq p_2$, i.e. p_1 and p_2 have at most $r-1$ prime icosians as factors in common. Note that, in this case

$$g := \mathrm{glcd}(p_1, p_2) = \mathrm{glcd}(p_1, \tfrac{|p_2|^2}{\sqrt{5}}) = \mathrm{glcd}(p_1, \tfrac{|p_1\widetilde{p}_1|}{\sqrt{5}}) = \mathrm{glcd}(p_1, \tfrac{|p_2\widetilde{p}_2|}{\sqrt{5}}),$$

since $|p_1\widetilde{p_1}| = ((\sqrt{5})^r(-\sqrt{5})^r)^{\frac{1}{2}} = (\sqrt{5})^r = |p_1|^2 = |p_2|^2$. We deduce that p_1 and p_2 have at least $r-1$ prime factors in common. Hence $|g|^2 = \frac{|p_1|^2}{\sqrt{5}} = |g\widetilde{g}| = (\sqrt{5})^{r-1}$ is not an integer, i.e. g is not admissible for L.

Define $g_1 = \mathrm{glcd}(p_1, \frac{|p_1|^2}{5})$, which is the greatest admissible divisor of g. Of course $|g_1|^2 = \frac{|p_1|^2}{5}$, by Corollary 1.25. So we have with (3.27)

$$\varphi_+(p_1\mathbb{I}) \subset \varphi_+(p_1\mathbb{I}) + \varphi_+(p_2\mathbb{I}) = \varphi_+(g\mathbb{I}) \subset \varphi_+(g_1\mathbb{I})$$

and $[\varphi_+(g_1\mathbb{I}) : \varphi_+(p_1\mathbb{I})] = \frac{[L:\varphi_+(p_1\mathbb{I})]}{[L:\varphi_+(g_1\mathbb{I})]} = \frac{|p_1|^2}{|g_1|^2} = 5$; see Theorem 3.37. Hence either $\varphi_+(p_1\mathbb{I}) = \varphi_+(p_1\mathbb{I}) + \varphi_+(p_2\mathbb{I})$ – in this case the proof is complete – or $\varphi_+(g\mathbb{I}) = \varphi_+(g_1\mathbb{I})$. We finish the proof by showing that the latter is impossible.

Assume the latter is the case, i.e.$[\varphi_+(g\mathbb{I}) : \varphi_+(p_1\mathbb{I})] = 5$. By Lemma 3.18 and Theorem 3.37 we see that

$$[L[\tau] : \varphi_+(g_1\mathbb{I}) + \tau\varphi_+(g_1\mathbb{I})] = [L[\tau] : L(g_1) + \tau L(g_1)] = \mathrm{nr}(g_1)^2 = \tfrac{\mathrm{nr}(p_1)^2}{25}$$

and $[\mathbb{I} : \varphi_+(g_1\mathbb{I}) + \tau\varphi_+(g_1\mathbb{I})] = [\mathbb{I} : L[\tau]]\Sigma_{L[\tau]}(g_1) = 5\,\mathrm{nr}(g_1)^2 = \frac{\mathrm{nr}(p_1)^2}{5}$. Furthermore, $\varphi_+(g_1\mathbb{I})+\tau\varphi_+(g_1\mathbb{I}) = \varphi_+(g\mathbb{I})+\tau\varphi_+(g\mathbb{I}) \subset g\mathbb{I}+\mathbb{I}\widetilde{g}$ and $[\mathbb{I} : g\mathbb{I}+\mathbb{I}\widetilde{g}] = \mathrm{N}(|g|^2) = \frac{\mathrm{nr}(p_1)^2}{5}$. Hence $g\mathbb{I} + \mathbb{I}\widetilde{g} = \varphi_+(g_1\mathbb{I}) + \tau\varphi_+(g_1\mathbb{I}) \subset L[\tau]$.

By Corollary 1.17 we know that $\sqrt{5}\mathbb{I} \subset L[\tau]$. Let $q := \mathrm{glcd}(g, \sqrt{5})$, i.e. $q\mathbb{I} \subset L[\tau]$. Since we know by Lemma 1.27 that $\widetilde{q} = \mathrm{grcd}(\widetilde{g}, -\sqrt{5})$, we infer that $\mathbb{I}\widetilde{q} \subset L[\tau]$. Proposition 1.16 tells us that for $r \in \mathbb{I}$, we have $qr - \widetilde{r}\widetilde{q} \in \sqrt{5}\mathbb{I}$, which implies that $\widetilde{r}\widetilde{q} = q(r-s)$ for $s \in \mathbb{I}$ and hence $\mathbb{I}\widetilde{q} \subset q\mathbb{I}$. As $[\mathbb{I} : q\mathbb{I}] = \mathrm{N}(\mathrm{nr}(q)^2) = [\mathbb{I} : \mathbb{I}\widetilde{q}]$, see Lemma 1.9, it follows that $q\mathbb{I} = \mathbb{I}\widetilde{q}$. As a divisor of g, the icosian q is primitive. This leads to a contradiction with Lemma 1.20. □

A combination of Theorem 3.40 and Lemma 3.41 leads to the following necessary and sufficient condition on the parameterising icosians for the remaining case.

THEOREM 3.42. *Let $p_1, p_2 \in \mathbb{I}$ be primitive, admissible and such that $|p_1|^2$ and $|p_2|^2$ are divisible by 5. Then*

$$L \cap \tfrac{1}{|p_1\widetilde{p_1}|} p_1 L\widetilde{p_1} = L \cap \tfrac{1}{|p_2\widetilde{p_2}|} p_2 L\widetilde{p_2}$$

if and only if

$$|p_1|^2 = |p_2|^2 \quad \textit{and} \quad \operatorname{glcd}(p_1, \tfrac{|p_1\widetilde{p_1}|}{\sqrt{5}}) = \operatorname{glcd}(p_2, \tfrac{|p_2\widetilde{p_2}|}{\sqrt{5}}).$$

PROOF. Since p_1 and p_2 are admissible, we know by Lemma 3.12 that in the prime factorisation of $|p_1|^2$ and $|p_2|^2$ every ramified prime number as well as every prime number that does not occur in $\gcd(|p_i|^2, |\widetilde{p_i}|^2)$, for $i = 1, 2$ respectively, has an even exponent. Here, we need to order the prime factors of $|p_1|^2$ and $|p_2|^2$. As this is done for p_1 and p_2 in exactly the same way, we set $i = 1, 2$. A particular ordering of the prime factors of $|p_i|^2$ with maximal exponents is

$$|p_i|^2 = (\sqrt{5})^{2r_i} \alpha_{i1}^{s_{i1}} \dots \alpha_{im_i}^{s_{im_i}} \alpha_{i(m_i+1)}^{t_{m_i+1}} \dots \alpha_{in_i}^{t_{n_i}} \alpha_{i1}^{s_{i1}} \dots \alpha_{im_i}^{s_{im_i}}, \tag{3.31}$$

where $\alpha_{i1}, \dots, \alpha_{im_i}$ are splitting primes which do not divide $\gcd(|p_i|^2, |\widetilde{p_i}|^2)$, and the remaining primes are denoted by $\alpha_{i(m_i+1)}, \dots, \alpha_{in_i}$. Define $s_i = s_{i1} + \dots + s_{im_i}$ and $t_i = t_{m_i+1} + \dots + t_{n_i}$. While leaving the prime factors of (3.31) in exactly the same order we rename them as

$$|p_i|^2 = (\sqrt{5})^{2r_i} \beta_{i1} \dots \beta_{i(2s_i+t_i)}.$$

By Theorem 1.23 there are prime icosians $u_{i1}, \dots, u_{i2r_i}$, $v_{i1}, \dots, v_{i2r_i}$, $a_{i1}, \dots, a_{i(2s_i+t_i)}$ and $b_{i1}, \dots, b_{i(2s_i+t_i)}$ such that

$$\begin{aligned} p_i &= u_{i1} \dots u_{i2r_i} a_{i1} \dots a_{i(2s_i+t_i)}, \text{ where } |u_{ij}|^2 = \sqrt{5} \text{ and } |a_{ij}|^2 = \beta_{ij}, \\ &= b_{i1} \dots b_{i(2s_i+t_i)} v_{i1} \dots v_{i2r_i}, \text{ where } |b_{ij}|^2 = \beta_{ij} \text{ and } |v_{ij}|^2 = \sqrt{5}. \end{aligned} \tag{3.32}$$

To simplify this notation we define $a_i = a_{i1} \dots a_{i(2s_i+t_i)}, b_i = b_{i1} \dots b_{i(2s_i+t_i)}$, $u_i = u_{i1} \dots u_{i2r_i}$ and $v_i = v_{i1} \dots v_{i2r_i}$. Hence $|u_i|^2 = (\sqrt{5})^{2r_i} = |v_i|^2$ and $\sqrt{5}$

does not divide $|a_i|^2 = |b_i|^2$. Note that

$$\begin{aligned} |a_i\widetilde{a_i}| &= \alpha_{i1}^{s_{i1}} \dots \alpha_{im_i}^{s_{im_i}} \alpha_{i(m_i+1)}^{t_{m_i+1}} \dots \alpha_{in_i}^{t_{n_i}} (\alpha')_{i1}^{s_{i1}} \dots (\alpha')_{im_i}^{s_{im_i}} = |b_i\widetilde{b_i}| \\ &= a_{i1} \dots a_{i(s_i+t_i)} \overline{a_{i1} \dots a_{i(s_i+t_i)}} (\alpha')_{i1}^{s_{i1}} \dots (\alpha')_{im_i}^{s_{im_i}} \qquad (3.33) \\ &= b_{i1} \dots b_{i(s_i+t_i)} \overline{b_{i1} \dots b_{i(s_i+t_i)}} (\alpha')_{i1}^{s_{i1}} \dots (\alpha')_{im_i}^{s_{im_i}}, \end{aligned}$$

which reveals that $\operatorname{glcd}(a_i, |a_i\widetilde{a_i}|) = a_{i1} \dots a_{i(s_i+t_i)}$ and $\operatorname{glcd}(b_i, |b_i\widetilde{b_i}|) = b_{i1} \dots b_{i(s_i+t_i)}$. Similarly, $\operatorname{glcd}(u_i, \frac{|u_i\widetilde{u_i}|}{\sqrt{5}}) = u_{i1} \dots u_{i(2r_i-1)}$ and $\operatorname{glcd}(v_i, \frac{|v_i\widetilde{v_i}|}{\sqrt{5}}) = v_{i1} \dots v_{i(2r_i-1)}$ since

$$\begin{aligned} |u_i\widetilde{u_i}| = |v_i\widetilde{v_i}| = (\sqrt{5})^{2r_i} &= u_{i1} \dots u_{i(2r_i)} \overline{u_{i1} \dots u_{i(2r_i)}} \\ &= v_{i1} \dots v_{i(2r_i)} \overline{v_{i1} \dots v_{i(2r_i)}}. \end{aligned}$$

Since $|a_i\widetilde{a_i}| = |b_i\widetilde{b_i}|$ and $|u_i\widetilde{u_i}| = |v_i\widetilde{v_i}|$ are clearly integers, a_i, b_i, u_i and v_i are admissible and as divisors of primitive icosians themselves primitive.

If $L \cap \frac{1}{|q_1\widetilde{p_1}|} p_1 L\widetilde{p_1} = L \cap \frac{1}{|p_2\widetilde{p_2}|} p_2 L\widetilde{p_2}$, Lemma 3.38 tells us that $\operatorname{den}(p_1) = \operatorname{den}(p_2)$ and $\Sigma(p_1) = \Sigma(p_2)$, hence by Proposition 3.10 we see that $|p_1\widetilde{p_1}| = |p_2\widetilde{p_2}| = (\sqrt{5})^{2r_1}|a_1\widetilde{a_1}| = (\sqrt{5})^{2r_2}|a_2\widetilde{a_2}|$, which implies that $r_1 = r_2 =: r$ and $|a_1\widetilde{a_1}| = |a_2\widetilde{a_2}|$. With Theorem 3.37 we conclude that

$$\begin{aligned} \operatorname{lcm}(|p_1|^2, |\widetilde{p}_1|^2) = \operatorname{lcm}(|p_2|^2, |\widetilde{p}_2|^2) &= (\sqrt{5})^r \operatorname{lcm}(|a_1|^2, |\widetilde{a_1}|^2) \\ &= (\sqrt{5})^r \operatorname{lcm}(|a_2|^2, |\widetilde{a_2}|^2) \end{aligned}$$

and furthermore $\Sigma(a_1) = \Sigma(a_2) = \Sigma(b_1) = \Sigma(b_2) =: m$, as well as, $\Sigma(u_1) = \Sigma(u_2) = \Sigma(v_1) = \Sigma(v_2) = (\sqrt{5})^r$. Since m and $(\sqrt{5})^r$ are relatively prime, we conclude with Lemma 3.5 that

$$\begin{aligned} m(L \cap \tfrac{1}{|u_1\widetilde{u_1}|} u_1 L\widetilde{u_1}) &= mL \cap (L \cap \tfrac{1}{|p_1\widetilde{p_1}|} p_1 L\widetilde{p_1}) \\ &= mL \cap (L \cap \tfrac{1}{|p_2\widetilde{p_2}|} p_2 L\widetilde{p_2}) = m(L \cap \tfrac{1}{|u_2\widetilde{u_2}|} u_2 L\widetilde{u_2}) \end{aligned}$$

as well as

$$(\sqrt{5})^r (L \cap \tfrac{1}{|b_1\widetilde{b_1}|} b_1 L\widetilde{b_1}) = (\sqrt{5})^r L \cap (L \cap \tfrac{1}{|p_1\widetilde{p_1}|} p_1 L\widetilde{p_1})$$

$$= (\sqrt{5})^r L \cap (L \cap \tfrac{1}{|p_2\widetilde{p_2}|} p_2 L\widetilde{p_2}) = (\sqrt{5})^r (L \cap \tfrac{1}{|b_2\widetilde{b_2}|} b_2 L\widetilde{b_2}).$$

Now, the application Lemma 3.41 and Theorem 3.40 reveals that $\operatorname{glcd}(u_1, \frac{|u_1\widetilde{u_1}|}{\sqrt{5}}) = \operatorname{glcd}(u_2, \frac{|u_2\widetilde{u_2}|}{\sqrt{5}})$, $|b_1|^2 = |b_2|^2$ and $\operatorname{glcd}(b_1, |b_1\widetilde{b_1}|) = \operatorname{glcd}(b_2, |b_2\widetilde{b_2}|)$, which implies $|p_1|^2 = |b_1|^2(\sqrt{5})^r = |b_2|^2(\sqrt{5})^r = |p_2|^2$. Furthermore, we conclude that

$$\operatorname{glcd}(p_1, (\sqrt{5})^{2r-1}) = \operatorname{glcd}(p_2, (\sqrt{5})^{2r-1}) \quad \text{and} \quad \operatorname{glcd}(p_1, |b_1\widetilde{b_1}|) = \operatorname{glcd}(p_2, |b_1\widetilde{b_1}|).$$

As $\gcd((\sqrt{5})^{2r-1}, |b_1\widetilde{b_1}|) = 1$ we infer with Lemma 1.26 that

$$\begin{aligned}
\operatorname{glcd}(p_1, \tfrac{|p_1\widetilde{p_1}|}{\sqrt{5}})\mathbb{I} &= \operatorname{glcd}(p_1, |b_1\widetilde{b_1}|(\sqrt{5})^{2r-1})\mathbb{I} \\
&= \operatorname{glcd}(p_1, |b_1\widetilde{b_1}|)\mathbb{I} \cap \operatorname{glcd}(p_1, (\sqrt{5})^{2r-1})\mathbb{I} \\
&= \operatorname{glcd}(p_2, |b_1\widetilde{b_1}|)\mathbb{I} \cap \operatorname{glcd}(p_2, (\sqrt{5})^{2r-1})\mathbb{I} \\
&= \operatorname{glcd}(p_2, |b_1\widetilde{b_1}|(\sqrt{5})^{2r-1})\mathbb{I} \\
&= \operatorname{glcd}(p_2, \tfrac{|p_2\widetilde{p_1}|}{\sqrt{5}})\mathbb{I}.
\end{aligned}$$

Conversely, if $|p_1|^2 = |p_2|^2$ and $\operatorname{glcd}(p_1, \frac{|p_1\widetilde{p_1}|}{\sqrt{5}}) = \operatorname{glcd}(p_2, \frac{|p_2\widetilde{p_2}|}{\sqrt{5}})$, we conclude again with Theorem 1.23 that $\operatorname{glcd}(u_1, \frac{|u_1\widetilde{u_1}|}{\sqrt{5}}) = \operatorname{glcd}(u_2, \frac{|u_2\widetilde{u_2}|}{\sqrt{5}})$ and $\operatorname{glcd}(b_1, |b_1\widetilde{b_1}|) = \operatorname{glcd}(b_2, |b_2\widetilde{b_2}|)$. Hence by Lemma 3.41 and Theorem 3.40 we know that

(3.34)

$$L \cap \tfrac{1}{|u_1\widetilde{u_1}|} u_1 L\widetilde{u_1} = L \cap \tfrac{1}{|u_2\widetilde{u_1}|} u_2 L\widetilde{u_2} \quad \text{and} \quad L \cap \tfrac{1}{|b_1\widetilde{b_1}|} b_1 L\widetilde{b_1} = L \cap \tfrac{1}{|b_2\widetilde{b_2}|} b_2 L\widetilde{b_2}.$$

Obviously, $\operatorname{lcm}(|p_1|^2, |\widetilde{p}_1|^2) = \operatorname{lcm}(|p_2|^2, |\widetilde{p}_2|^2) = (\sqrt{5})^r \operatorname{lcm}(|a_1|^2, |\widetilde{a_1}|^2) = (\sqrt{5})^r \operatorname{lcm}(|a_2|^2, |\widetilde{a_2}|^2)$. Hence Theorem 3.37 implies that $\Sigma(a_1) = \Sigma(a_2) = \Sigma(b_1) = \Sigma(b_2) =: m$, as well as, $\Sigma(u_1) = \Sigma(u_2) = \Sigma(v_1) = \Sigma(v_2) = (\sqrt{5})^r$. Since m and $(\sqrt{5})^r$ are relatively prime, we conclude with Corollary 3.4, for

$i = 1, 2$ the following:

$$\begin{aligned} L \cap R(p_i)L &= L \cap R(b_i)L \cap R(b_i)R(v_i)L = (L \cap R(b_i)L) \cap R(b_i)(L \cap R(v_i)L) \\ &= L \cap R(u_i)L \cap R(u_i)R(a_i)L = (L \cap R(u_i)L) \cap R(u_i)(L \cap R(a_i)L) \\ &\subset (L \cap R(b_i)L) \cap (L \cap R(u_i)L) \end{aligned}$$

The index of $(L \cap R(b_i)L) \cap (L \cap R(u_i)L)$ in L, is on the one hand a divisor of $\Sigma(p_i) = m(\sqrt{5})^r = \Sigma(b_i)\Sigma(u_i)$, and on the other hand a multiple of $\Sigma(b_i)$ and $\Sigma(u_i)$, which are relatively prime. So we know that

$$[L : (L \cap R(b_i)L) \cap (L \cap R(u_i)L)] = \Sigma(p_i)$$

and hence $(L \cap R(b_i)L) \cap (L \cap R(u_i)L) = L \cap R(p_i)L$. Now, we see with (3.34) that $L \cap R(p_1)L = L \cap R(p_2)L$. □

3.4.1. Generating Functions. For the derivation of the generating functions $\Phi_L^{\mathrm{rot}}(s)$ and $\Phi_L(s)$ from (3.3) and (3.4) we need the following

Proposition 3.43. *Let $f_K^{\mathsf{pr}}(\alpha)$ be the number of primitive right ideals in $\mathbb{I}$ with K-index $\alpha^2 \in \mathbb{Z}[\tau]$. Then, the arithmetic function $f_K^{\mathsf{pr}}(\alpha)$ is multiplicative and given by*

$$f_K^{\mathsf{pr}}(\alpha) \;=\; \begin{cases} 1, & \text{if } \alpha = 1, \\ 6 \cdot 5^{r-1}, & \text{if } \alpha = \sqrt{5}^r, \\ (p^2+1)p^{2r-2}, & \text{if } \alpha = p^r \text{ and } p \equiv \pm 2 \ (5), \\ (p+1)p^{r-1}, & \text{if } \alpha = \pi^r, p = \pi\pi' \text{ and } p \equiv \pm 1 \ (5). \end{cases}$$

Proof. By Lemma 1.8 we know that $[\mathbb{I} : q\mathbb{I}]_K = |q|^4$. Hence the multiplicativity of f_K^{pr} is inherited from the unique factorisation in $\mathbb{I}$ and, as for every multiplicative function, $f_K^{\mathsf{pr}}(1) = 1$. Furthermore, we know by Lemma 1.7 that for every right ideal $q\mathbb{I}$

$$[\mathbb{I} : q\mathbb{I}] = \mathrm{N}([\mathbb{I} : q\mathbb{I}]_K) = \mathrm{N}(|q|^4).$$

Since $\sqrt{5}$ is a ramified prime in $\mathbb{Z}[\tau]$, this means that every primitive right ideal $q\mathbb{I}$ with coset-counting index $[\mathbb{I} : q\mathbb{I}] = \mathrm{N}(|q|^4) = 5^{2r}$, has a K-index of $[\mathbb{I} : q\mathbb{I}]_K = |q|^4 = \sqrt{5}^{2r}$. Hence $f_K^{\mathsf{pr}}(\sqrt{5}^r) = f^{\mathsf{pr}}(5^r) = 6 \cdot 5^{r-1}$, as defined in (2.8), which gives the number of primitive right ideals in $\mathbb{I}$ of coset-counting index 5^{2r}.

Similarly, every primitive right ideal $q\mathbb{I}$ with coset-counting index $[\mathbb{I} : q\mathbb{I}] = \mathrm{N}(|q|^4) = p^{2r}$, where $p \equiv \pm 2\ (5)$, has a K-index of $[\mathbb{I} : q\mathbb{I}]_K = |q|^4 = p^r$, since p is an inert prime of $\mathbb{Z}[\tau]$. Note that r is even, as already $\mathrm{N}(|q|^2) = p^r$ is a square in $\mathbb{N}$. So we have in this case $f_K^{\mathsf{pr}}(p^{\frac{r}{2}}) = f^{\mathsf{pr}}(p^r)$ or equivalently, with arbitrary $r \in \mathbb{N}$, $f_K^{\mathsf{pr}}(p^r) = f^{\mathsf{pr}}(p^{2r}) = p^{2r} + p^{2r-2}$.

The remaining case with $p \equiv \pm 1(5)$, where p splits as $p = \pi\pi'$, has to be treated differently. Observe, from the expansion of $\zeta_{\mathbb{I}}^{\mathsf{pr}}(s)$ as an Euler product in (2.7), that

$$\frac{1 + p^{-2s}}{1 - p^{1-2s}} \;=\; 1 + \sum_{r \geq 1} (p^r + p^{r-1}) p^{-2rs},$$

happens to be the generating function for the primitive right ideals $q\mathbb{I}$ of p-power index such that $|q|^4$ is a power of π. Those with $|q|^4$ a power of π' produce the same Euler factor. Hence we infer that $f_K^{\mathsf{pr}}(\pi^r) = p^r + p^{r-1}$. $\square$

We start with the derivation of $\Phi_L^{\mathrm{rot}}(s)$ from (3.4). Recall from (3.8) that every coincidence rotation of the lattice L is parameterised by a primitive admissible icosian q. By Theorem 3.37 its index is $\Sigma(q) = \mathrm{lcm}(\mathrm{nr}(q), \mathrm{nr}(\widetilde{q}))$. Note that an icosian q with $\mathrm{nr}(q) = 1$ is necessarily admissible and primitive. Moreover, $1 = \mathrm{nr}(q) = \mathrm{nr}(\widetilde{q}) = \mathrm{lcm}(\mathrm{nr}(q), \mathrm{nr}(\widetilde{q})) = \Sigma(q)$, so there are $120 = |I|$, see (1.21), rotations in $\mathrm{SOC}(L)$ with coincidence index 1. Therefore, let $120 g_{\mathrm{rot}}(m)$ be the number of coincidence rotations of the lattice L of index m.

Clearly, we have $g_{\mathrm{rot}}(1) = 1$. For $m = 5^r$ we realise that

$$\Sigma(q) = \mathrm{lcm}(\mathrm{nr}(q), \mathrm{nr}(\widetilde{q})) = \sqrt{5}^{2r} = 5^r = |q|^2 = |\widetilde{q}|^2 = |q\widetilde{q}| \tag{3.35}$$

and $\mathrm{nr}(q)^2 = [\mathbb{I} : q\mathbb{I}]_K$. Hence $g_{\mathrm{rot}}(5^r) = f_K^{\mathsf{pr}}(5^r) = 6 \cdot 5^{r-1}$. Similarly, for $m = p^r$ with $p \equiv \pm 2 \mod 5$ we observe that $\Sigma(q) = p^r = \mathrm{nr}(q) = \mathrm{nr}(\widetilde{q}) = |q\widetilde{q}|$ and hence $g_{\mathrm{rot}}(p^r) = f_K^{\mathsf{pr}}(p^r) = p^{2r} + p^{2r+2}$.

For the remaining case $m = p^r$ with $p \equiv \pm 1 \mod 5$, where p splits as $p = \pi\pi'$ in $\mathbb{Z}[\tau]$, the derivation is slightly more complicated, because one has to keep track of how the algebraically conjugate primes of $\mathbb{Z}[\tau]$ are distributed in $\mathrm{nr}(q)$ and $\mathrm{nr}(\widetilde{q})$. Note that there are $s, t \in \mathbb{N}$ such that $\mathrm{nr}(q) = \pi^s(\pi')^t$ and

$$\Sigma(q) = \mathrm{lcm}(\mathrm{nr}(q), \mathrm{nr}(\widetilde{q})) = \mathrm{lcm}(\mathrm{nr}(q), \mathrm{nr}(q)') = p^{\max(s,t)} = p^r. \tag{3.36}$$

Moreover, q is admissible if and only if $s + t$ is even. In order to determine $g_{\mathrm{rot}}(p^r)$ we have to sum up the number of all primitive quaternions q with $\mathrm{nr}(q) = \pi^s(\pi')^t$, such that $s + t$ is even. This gives with Proposition 3.43:

$$\begin{aligned} g_{\mathrm{rot}}(p^r) &= f_K^{\mathsf{pr}}(\pi^r) f_K^{\mathsf{pr}}(\pi'^r) + \sum_{\ell=1}^{\lfloor \frac{r}{2} \rfloor} f_K^{\mathsf{pr}}(\pi^r) f_K^{\mathsf{pr}}(\pi'^{(r-2\ell)}) + \sum_{\ell=1}^{\lfloor \frac{r}{2} \rfloor} f_K^{\mathsf{pr}}(\pi^{r-2\ell}) f_K^{\mathsf{pr}}(\pi'^r) \\ &= (p+1)^2 p^{2r-2} + 2(p+1)p^{r-1} \sum_{\ell=1}^{\lfloor \frac{r}{2} \rfloor} f_K^{\mathsf{pr}}(\pi^{r-2\ell}) \end{aligned}$$

For odd r we have

$$\begin{aligned} g_{\mathrm{rot}}(p^r) &= (p+1)^2 p^{2r-2} + 2(p+1)p^{r-1} \sum_{\ell=1}^{\frac{r-1}{2}} (p+1)p^{r-2\ell-1} \\ &= (p+1)^2 p^{2r-2} + 2(p+1)^2 p^{r-1} \frac{(p^{r-1}-1)}{p^2-1} \\ &= \frac{(p+1)}{(p-1)} p^{r-1}(p^{r+1} + p^{r-1} - 2). \end{aligned}$$

The calculation for even r is slightly different but leads to the same result:

$$\begin{aligned} g_{\mathrm{rot}}(p^r) &= (p+1)^2 p^{2r-2} + 2(p+1)p^{r-1} \sum_{\ell=1}^{(r/2)-1} (p+1)p^{r-2\ell-1} + 2(p+1)p^{r-1} \\ &= (p+1)^2 p^{2r-2} + 2(p+1)^2 p^{r-1} \frac{(p^{r-1}-p)}{p^2-1} + 2(p+1)p^{r-1} \\ &= \frac{(p+1)}{(p-1)} p^{r-1}(p^{r+1} + p^{r-1} - 2). \end{aligned}$$

In summary g_{rot} is thus given by

$$g_{\mathrm{rot}}(p^r) = \begin{cases} 6 \cdot 5^{2r-1}, & \text{if } p = 5\,, \\ \frac{p+1}{p-1}\, p^{r-1}(p^{r+1} + p^{r-1} - 2), & \text{if } p \equiv \pm 1 \ (5)\,, \\ (p^2+1)p^{2r-2}, & \text{if } p \equiv \pm 2 \ (5)\,. \end{cases} \tag{3.37}$$

Note that g_{rot} is multiplicative due to Theorem 1.23 together with the multiplicativity of the coincidence index; see Theorem 3.3. Hence (3.37) fixes $g_{\mathrm{rot}}(m)$ for all $m \in \mathbb{N}$, and this completes the derivation of $\Phi_L^{\mathrm{rot}}(s) = \sum_{m=1}^{\infty} \frac{g_{\mathrm{rot}}(m)}{m^s}$.

Based on (3.37) we calculate now $g(m)$, the number of CSLs of L of index m, and hence derive $\Phi_L(s)$. We have to keep in mind that non-equivalent coincidence rotations may lead to the same CSL.

We start again with the case $m = 5^r$. By Lemma 3.41 two primitive admissible icosians q_1 and q_2 with $\Sigma(q_1) = \Sigma(q_2) = |q_1|^2 = |q_2|^2 = 5^r$, compare (3.35), generate the same CSL if and only if $\mathrm{glcd}(q_1, \frac{|q_1\widetilde{q}_1|}{\sqrt{5}}) = \mathrm{glcd}(q_2, \frac{|q_2\widetilde{q}_2|}{\sqrt{5}}) = \mathrm{glcd}(q_1, 5^{(r-\frac{1}{2})}) = \mathrm{glcd}(q_2, 5^{(r-\frac{1}{2})})$ or equivalently if and only if q_1 and q_2 differ at most in their rightmost prime factor. Thus the CSL is uniquely determined by the first $2r-1$ prime factors of q_1 and q_2 of norm $\sqrt{5}$. Since primitive icosians with $\Sigma(q) = |q|^2 = |q\widetilde{q}| = 5^r$ are necessarily admissible this means that $g(5^r) = f_K^{\mathsf{pr}}(\sqrt{5}^{(2r-1)}) = 6 \cdot 5^{2r-2}$.

For $m = p^r$ with $p \equiv \pm 2 \mod 5$ Theorem 3.40 implies that two primitive admissible icosians q_1 and q_2 with $\Sigma(q_1) = \Sigma(q_2) = |q_1|^2 = |q_2|^2 = p^r$ generate the same CSL if and only if $\mathrm{glcd}(q_1, |q_1\widetilde{q}_1|) = \mathrm{glcd}(q_2, |q_2\widetilde{q}_2|)$. Since $|q_1\widetilde{q}_1| = |q_1|^2 = |q_2|^2 = |q_2\widetilde{q}_2|$ the latter just means that there is a $u \in \mathbb{I}^\times$ such that $q_1 = q_2 u$. Hence $g(p^r) = f_K^{\mathsf{pf}}(p^r) = (p^2+1)p^{2r-2}$.

Again the remaining case $m = p^r$ with $p \equiv \pm 1 \mod 5$, where p splits as $p = \pi\pi'$ in $\mathbb{Z}[\tau]$, is slightly more complicated. By Theorem 3.40 two primitive admissible quaternions q_1 and q_2 with $\Sigma(q_1) = \Sigma(q_2) = p^r$ generate the same CSL if and only if $|q_1|^2 = |q_2|^2$ and $\mathrm{glcd}(q_1, |q_1\widetilde{q}_1|) = \mathrm{glcd}(q_2, |q_2\widetilde{q}_2|)$. Let $s, t \in \mathbb{N}$ such that $|q_1|^2 = |q_2|^2 = \pi^s(\pi')^t$, compare (3.36), and recall that the

admissibility requires $s+t$ to be even. With this notation we have $|q_1\widetilde{q}_1| = p^{\frac{s+t}{2}} = |q_2\widetilde{q}_2|$. Obviously, the condition that $\mathrm{glcd}(q_1, p^{\frac{s+t}{2}}) = \mathrm{glcd}(q_2, p^{\frac{s+t}{2}})$ is equivalent to $\mathrm{glcd}(q_1, \pi^{\frac{s+t}{2}}) = \mathrm{glcd}(q_2, \pi^{\frac{s+t}{2}})$ and $\mathrm{glcd}(q_1, \pi'^{\frac{s+t}{2}}) = \mathrm{glcd}(q_2, \pi'^{\frac{s+t}{2}})$. This means that in the prime decomposition of q_1 and q_2 the first $\frac{s+t}{2}$ prime factors from the left of norm π and π', respectively, must be the same whereas the rest may differ. In other words the CSL is uniquely determined by the first $\frac{s+t}{2}$ prime factors of norm π and π'. Since at least one of the inequalities $s \leq \frac{s+t}{2}$ and $t \leq \frac{s+t}{2}$ holds, this gives with Proposition 3.43:

$$\begin{aligned} g(p^r) &= f_K^{\mathrm{pr}}(\pi^r)^2 + 2\sum_{\ell=1}^{\lfloor r/2 \rfloor} f_K^{\mathrm{pr}}(\pi^{r-\ell}) f_K^{\mathrm{pr}}(\pi^{r-2\ell}) \\ &= (p+1)^2 p^{2r-2} + 2(p+1)\sum_{\ell=1}^{\lfloor r/2 \rfloor} p^{r-\ell-1} f_K^{\mathrm{pr}}(\pi^{r-2\ell}). \end{aligned}$$

If r is odd we get:

$$\begin{aligned} g(p^r) &= (p+1)^2 p^{2r-2} + 2(p+1)^2 \sum_{\ell=1}^{(r-1)/2} p^{2r-3\ell-2} \\ &= (p+1)^2 p^{2r-2} + 2(p+1)^2 \frac{p^{(2r-2)} - p^{\frac{(r-1)}{2}}}{p^3-1} \\ &= \frac{(p+1)^2}{p^3-1}\left(p^{2r+1} + p^{2r+2} - 2p^{\frac{(r-1)}{2}}\right) \end{aligned}$$

If r is even we get:

$$\begin{aligned} g(p^r) &= (p+1)^2 p^{2r-2} + 2(p+1)^2 \sum_{\ell=1}^{(r-2)/2} p^{2r-3\ell-2} + 2(p+1)p^{r/2-1} \\ &= (p+1)^2 p^{2r-2} + 2(p+1)^2 \frac{p^{(2r-2)} - p^{\frac{(r+2)}{2}}}{p^3-1} + 2(p+1)p^{r/2-1} \\ &= \frac{(p+1)^2}{p^3-1}\left(p^{2r+1} + p^{2r+2} - 2p^{r/2-1}\frac{(p^2+1)}{p+1}\right) \end{aligned}$$

In summary g is thus given by

(3.38)
$$g(p^r) \;=\; \begin{cases} 6\cdot 5^{2r-2}, & \text{if } p=5\,,\\ \frac{(p+1)^2}{p^3-1}\left(p^{2r+1}+p^{2r+2}-2p^{\frac{(r-1)}{2}}\right) & \text{if } r \text{ is odd, } p\equiv\pm 1\ (5)\,,\\ \frac{(p+1)^2}{p^3-1}\left(p^{2r+1}+p^{2r+2}-2p^{r/2-1}\frac{(p^2+1)}{p+1}\right), & \text{if } r \text{ is even, } p\equiv\pm 1\ (5)\,,\\ (p^2+1)p^{2r-2}, & \text{if } p\equiv\pm 2\ (5)\,. \end{cases}$$

Again g is multiplicative due to the unique factorisation in $\mathbb{I}$ together with the multiplicativity of the coincidence index, see Theorem 3.3. Hence (3.38) defines $g(m)$ for all $m\in\mathbb{N}$, and this completes the derivation of $\Phi_L(s) = \sum_{m=1}^{\infty}\frac{g(m)}{m^s}$.

The explicit expressions of the multiplicative arithmetic functions g_{rot} and g, from (3.37) and (3.38) respectively, lead by straightforward calculations involving geometric series to the following Euler product expansions of the corresponding Dirichlet series generating functions. As the lattices L and A_4 are isomorphic they clearly coincide. Note that the following theorem was already stated without proof in [**15**].

THEOREM 3.44. *Let $120\,g_{\mathrm{rot}}(m)$ be the number of coincidence rotations of A_4 of index m and $120\,g(m)$ be the number of CSLs of A_4 of index m. Then, $g_{\mathrm{rot}}(m)$ and $g(m)$ are multiplicative arithmetic functions, with Dirichlet series generating functions*

$$\begin{aligned}\Phi^{\mathrm{rot}}_{A_4}(s) \;&=\; \sum_{m=1}^{\infty}\frac{g_{\mathrm{rot}}(m)}{m^s} \;=\; \frac{\zeta_K(s-1)}{1+5^{-s}}\,\frac{\zeta(s)\,\zeta(s-2)}{\zeta(2s)\,\zeta(2s-2)}\\ &=\; \frac{1+5^{1-s}}{1-5^{2-s}}\prod_{p\equiv\pm1\,(5)}\frac{(1+p^{-s})\,(1+p^{1-s})}{(1-p^{1-s})\,(1-p^{2-s})}\prod_{p\equiv\pm2\,(5)}\frac{1+p^{-s}}{1-p^{2-s}}\\ &=\; 1+\frac{5}{2^s}+\frac{10}{3^s}+\frac{20}{4^s}+\frac{30}{5^s}+\frac{50}{6^s}+\frac{50}{7^s}+\frac{80}{8^s}+\frac{90}{9^s}+\frac{150}{10^s}+\frac{144}{11^s}+\dots\end{aligned}$$

and

$$\begin{aligned}\Phi_{A_4}(s) &= \sum_{m=1}^{\infty} \frac{g(m)}{m^s} \\ &= 1 + \tfrac{6\cdot 5^{-s}}{1-5^{2-s}} \prod_{p\equiv\pm 1\,(5)} \tfrac{(1+p^{-s}+2p^{1-s}+2p^{-2s}+p^{1-2s}+p^{1-3s})}{(1-p^{2-s})(1-p^{1-2s})} \prod_{p\equiv\pm 2\,(5)} \tfrac{1+p^{-s}}{1-p^{2-s}} \\ &= 1 + \frac{5}{2^s} + \frac{10}{3^s} + \frac{20}{4^s} + \frac{6}{5^s} + \frac{50}{6^s} + \frac{50}{7^s} + \frac{80}{8^s} + \frac{90}{9^s} + \frac{30}{10^s} + \frac{144}{11^s} + \dots\end{aligned}$$

where $\zeta(s)$ is Riemann's zeta function and $\zeta_K(s)$ denotes the Dedekind zeta function of the quadratic field $K = \mathbb{Q}(\sqrt{5}\,)$, see (1.16). □

Observe that $g_{\mathrm{rot}}(m) \geq g(m) > 0$ for all $m \in \mathbb{N}$. Since each element of $\mathrm{OC}(A_4)$ can be written as a product of a rotation with a reflection that maps A_4 onto itself we have

COROLLARY 3.45. *The coincidence spectrum of the root lattice A_4 is*

$$\Sigma(\mathrm{OC}(A_4)) = \Sigma(\mathrm{SOC}(A_4)) = \mathbb{N}.$$

□

We conclude this chapter with the analysis of the analytic properties of $\Phi_{A_4}^{\mathrm{rot}}(s)$. Recall from (2.9) the primitive Dirichlet character χ and its L-series $L(s,\chi) = \sum_{m=1}^{\infty} \chi(m) m^{-s}$, which is an entire function on the complex plane. Note that $\zeta_K(s) = \zeta(s)\, L(s,\chi)$, see [**87**, Chapter 11, Eq. (10)] for details. This means that the Dirichlet series $\Phi_{A_4}^{\mathrm{rot}}$ is related to zeta functions with well-defined analytic behaviour, namely

$$\Phi_{A_4}^{\mathrm{rot}}(s) = \frac{\zeta(s-1)\, L(s-1,\chi)}{1+5^{-s}} \, \frac{\zeta(s)\,\zeta(s-2)}{\zeta(2s)\,\zeta(2s-2)}.$$

Hence we can easily derive its analytic properties. $\Phi_{A_4}^{\mathrm{rot}}(s)$ is analytic in the open right half-plane $\{s = \sigma + it \mid \sigma > 3\}$, and has a simple pole at $s = 3$, due to the fact that $L(s,\chi)$ is analytic everywhere and $\zeta(s)$ is analytic except for a simple pole at $s = 1$ with residue 1, compare [**2**, Theorem 12.5]. The

corresponding residue is given by

$$\varrho := \operatorname{res}_{s=3} \Phi_{A_4}^{\mathrm{rot}}(s) \,=\, \frac{125}{126}\,\frac{\zeta_K(2)\,\zeta(3)}{\zeta(4)\,\zeta(6)} \,=\, \frac{450\sqrt{5}}{\pi^6}\,\zeta(3) \,\approx\, 1.258\,124\,, \tag{3.39}$$

which is based on the special values

$$\zeta(4) \,=\, \frac{\pi^4}{90}\,, \quad \zeta(6) \,=\, \frac{\pi^6}{945}\,, \quad \zeta_K(2) \,=\, \frac{2\pi^4}{75\sqrt{5}}\,,$$

together with $\zeta(3) \approx 1.202\,057$, compare [**12**] and [**2**, Theorem 12.17]. The value of $\zeta(3)$ is known to be irrational, but has to be calculated numerically.

With this information one can derive the asymptotic growth of $g_{\mathrm{rot}}(m)$ according to [**12**, Appendix]. Since the value of the arithmetic function $g_{\mathrm{rot}}(m)$ fluctuates heavily, this is done via the corresponding summatory function. One obtains, as $x \to \infty$, that

$$G(x) \,:=\, \sum_{m \leq x} g_{\mathrm{rot}}(m) \,\sim\, \varrho\,\frac{x^3}{3} \,\approx\, 0.419\,375\,x^3.$$

Clearly, this is also an upper bound for the asymptotic behaviour of $g(m)$ which counts the CSLs of A_4. We have not found a way to write the corresponding Dirichlet series Φ_{A_4} as a product of functions with well defined analytic behaviour. Hence its analytic properties and the asymptotic behaviour of $g(m)$ remain to be be analysed.

3.5. Related Results

In one dimension, the CSL problem becomes trivial, so that $\Phi(s) \equiv 1$ in this case. In two dimensions, a rather general approach to lattices and modules is possible via classic algebraic number theory, see for example [**71, 6**]. For the root lattice A_2, the CSL generating function reads

$$\Phi_{A_2}(s) \,=\, \Phi_{A_2}^{\mathrm{rot}}(s) \,=\, \prod_{p \equiv 1 \ (3)} \frac{1+p^{-s}}{1-p^{-s}} \,=\, \frac{1}{1+3^{-s}}\,\frac{\zeta_{\mathbb{Q}(\xi_3)}(s)}{\zeta(2s)}\,, \tag{3.40}$$

where $\xi_3 = e^{2\pi i/3}$. Here, the equality of $\Phi_{A_2}(s)$ and $\Phi_{A_2}^{\mathrm{rot}}(s)$ is a consequence of the commutativity of $\mathrm{SOC}(A_2)$. The simple coincidence spectrum of

this lattice is the multiplicative monoid of integers that is generated by the rational primes $p \equiv 1 \bmod 3$.

In three dimensions, various examples are derived [**3**] and were proved by quaternionic methods [**13**] similar to the ones used here. Among these cases is the root lattice A_3, which happens to be the face centred cubic lattice in 3-space, with generating function

$$\Phi_{A_3}(s) \,=\, \Phi_{A_3}^{\mathrm{rot}}(s) \,=\, \prod_{p\neq 2} \frac{1+p^{-s}}{1-p^{1-s}} \,=\, \frac{1-2^{1-s}}{1+2^{-s}}\,\frac{\zeta(s)\,\zeta(s-1)}{\zeta(2s)}\,. \tag{3.41}$$

The equality of the two Dirichlet series to the left is non-trivial, and was proved in [**13**] with an argument involving Eichler orders. The same formula also applies to the other cubic lattices in 3-space [**3**]. The simple coincidence spectrum is thus the set of odd integers, which is again a monoid.

Several of these results are also included by now in [**78**]. In 4-space, various other lattices and modules of interest exist, for which some results are given in [**3, 88**]. Beyond dimension 4, very little is known [**92, 91, 42**], though it should be possible to derive the simple coincidence spectra for certain classes of lattices.

Part II

Entropy of Powerfree Words

CHAPTER 4

Powerfree Words

The second part of this thesis considers words over finite alphabets that avoid certain repetitions in their sequence of letters. Another way to say this is that we consider powerfree words over finite alphabets. In particular, we will derive upper and lower bounds for the entropy of certain sets of powerfree words over finite alphabets.

This chapter, which is based on [**37**], introduces a new notation and terminology that is almost complementary to the one used in Part I. Moreover, the characterisation of certain classes of powerfree morphisms is reviewed and used to develop methods for the derivation of lower bounds for the entropy of certain sets of words. Furthermore, two methods to derive upper bounds for the entropy are described. In order to give the reader an idea to what kind of sets these methods will be applied to, the sets which are explicitly considered in Chapter 7, are already introduced here.

4.1. Notation and Definitions

4.1.1. Words. We define an *alphabet* $\mathbf{A}$ as a finite non-empty set of symbols called letters. For definiteness, we consider ℓ-letter alphabets

$$\mathbf{A}_\ell := \{0, 1, \ldots, \ell - 1\}\,.$$

Finite or infinite sequences of elements from $\mathbf{A}$ are called *words*. The *empty word* is denoted by ε while $w = w_1 \ldots w_n$, with $w_i \in \mathbf{A}$, stands for a finite word over $\mathbf{A}$ of length $|w| = n$. The length of the empty word is $|\varepsilon| := 0$. Note that we continue to use the notation $|S|$ for the cardinality of a set S as it will be clear from the context what is meant. A *subword* or *factor* of a word $w = w_1 \ldots w_n$ is defined as $w[i : j] := w_i \ldots w_j$, where $1 \leq i \leq j \leq n$.

If $i = j$, we write $w[i] = w_i$. For $1 \leq i \leq n$ the factor $w[1 : i]$ is called a *prefix* and the factor $w[i : n]$ is called a *suffix* of w.

The set of all finite words, the operation of concatenation of words and ε form the free monoid $\mathbf{A}^*$. For any subset $S \subset \mathbf{A}^*$ and any word $v \in \mathbf{A}^*$ we define

$$S(n) := \{w \in S \mid |w| = n\} \tag{4.1}$$

$$S^{(v)} := \{w \in S \mid v \text{ is a suffix of } w\} \tag{4.2}$$

$$\mathrm{Fact}(S) := \{v \in \mathbf{A}^* \mid v \text{ occurs as a factor of some } w \in S\}\,. \tag{4.3}$$

4.1.2. Morphisms. Let $\mathbf{A}$ and $\mathbf{B}$ be alphabets. A map $\varrho : \mathbf{A}^* \to \mathbf{B}^*$ is called a *morphism* if

$$\varrho(uv) = \varrho(u)\varrho(v) \text{ for all } u, v \in \mathbf{A}^*.$$

Obviously, a morphism ϱ is completely determined by $\varrho(a)$ for $a \in \mathbf{A}$ and satisfies $\varrho(\varepsilon) = \varepsilon$. It is called *$n$-uniform* or just *uniform*, if $|\varrho(a)| = n$ for all $a \in \mathbf{A}$.

A *permutation of letters* on an alphabet $\mathbf{A}_\ell$ is a bijective 1-uniform morphism
$\varrho : \mathbf{A}_\ell \to \mathbf{A}_\ell$. Two words $u, v \in \mathbf{A}_\ell^*$ are called *isomorphic*, if there exists a permutation of letters ϱ such that $\varrho(u) = v$.

4.1.3. Powerfreeness. An integer $p \in \mathbb{N}$ is called a *period* of

$$w = w_1 \dots w_n \in \mathbf{A}^*, \text{ if } w_i = w_{i+p} \text{ for all } i \in \{1, \dots, n-p\}\,.$$

The minimal period of w is denoted by $\mathrm{per}(w)$ and the ratio $\frac{|w|}{\mathrm{per}(w)}$ is called the *exponent* of w.

For a word w we define $w^0 := \varepsilon$, $w^1 := w$ and, for an integer $k > 1$, the power w^k as the concatenation of k occurrences of the word w. If $w \neq \varepsilon$,

w^k is called a *k-power*. Moreover, if $\mathrm{per}(w^k) = |w|$, i.e. w^k does not contain any shorter k-powers as factors, the exponent of w^k is $\frac{|w^k|}{\mathrm{per}(w)} = k$.

We say that a word u contains a k-power if at least one of its factors is a k-power. If u does not contain any k-power, it is called *k-powerfree* and we say that it avoids k-powers. By definition, the empty word ε is k-powerfree for all k. 2-powerfree and 3-powerfree words are called *squarefree* and *cubefree*, respectively.

Let $\mathbf{A}$ be an alphabet. We denote the set of k-powerfree words over $\mathbf{A}$ by

$$\mathcal{F}^{(k)}(\mathbf{A}) \subset \mathbf{A}^*. \tag{4.4}$$

Let $\alpha \in \mathbb{Q}$ with $\alpha > 1$. A non-empty word $w \in \mathbf{A}^*$ is called an *α-power* if there is a word $u \in \mathbf{A}^*$ which has a prefix v such that

$$w = u^k v \quad \text{and} \quad k + \frac{|v|}{|u|} = \alpha.$$

If $\mathrm{per}(w) = |u|$, i.e. if w does not contain any shorter k-powers as factors, the exponent of w is $\frac{|w|}{\mathrm{per}(w)} = \alpha$. For example 0123012 is a $\frac{7}{4}$-power of period 4.

For $\alpha \in \mathbb{R}$ we say that a word w is *α^+-powerfree (α-powerfree)* if it contains no β-power for any rational $\beta > \alpha$ ($\beta \geq \alpha$). 2^+-powerfree words are also called *overlapfree*, since they avoid factors of the form $auaua$, where $a \in \mathbf{A}$ and $u \in \mathbf{A}^*$. The set of all α^+-powerfree words over $\mathbf{A}$ is denoted by

$$\mathcal{F}^{(>\alpha)}(\mathbf{A}) \subset \mathbf{A}^*. \tag{4.5}$$

Note that we have the following proper inclusions

$$\mathcal{F}^{(\alpha)}(\mathbf{A}) \subset \mathcal{F}^{(>\alpha)}(\mathbf{A}) \subset \mathcal{F}^{(\alpha+1)}(\mathbf{A}).$$

Since powerfreeness only depends on the structure of a word and not on the actual letters, it suffices to consider equivalence classes up to permutations of letters. For many cases, including the examples we consider in Chapter 7, the equivalence classes can be represented by words with certain sufficiently long prefixes. For example, for $\mathcal{F} = \mathcal{F}^{(2)}(\mathbf{A}_3)$, the set of equivalence classes $\mathcal{F}'(m)$ can be represented by all ternary squarefree words of length m which start with 01. For $w \in \mathcal{F}$, where $\mathcal{F}$ stands for $\mathcal{F}^{(k)}(\mathbf{A})$ or $\mathcal{F}^{(>\alpha)}(\mathbf{A})$, we denote by w' its equivalence class and define

$$\mathcal{F}'(m) := \{w' \mid w \in \mathcal{F}(m)\}. \tag{4.6}$$

4.2. Characterisation of Integer Powerfree Morphisms

A morphism $\varrho : \mathbf{A}^* \to \mathbf{B}^*$, where $\mathbf{A}$ and $\mathbf{B}$ are alphabets, is called *k-powerfree*, if it maps k-powerfree words to k-powerfree words, i.e. if $\varrho(u)$ is k-powerfree for every k-powerfree word $u \in \mathbf{A}^*$. We express this symbolically as

$$\varrho(\mathcal{F}^{(k)}(\mathbf{A})) \subset \mathcal{F}^{(k)}(\mathbf{B}).$$

A set $T \subset \mathbf{A}^*$ is a *test-set* for the k-powerfreeness of (uniform) morphisms on $\mathbf{A}$, if for every (uniform) morphism $\varrho : \mathbf{A} \to \mathbf{B}$ the following holds: ϱ is k-powerfree if and only if $\varrho(T) \subset \mathcal{F}^{(k)}(\mathbf{B})$.

4.2.1. Squarefree Morphisms. A sufficient (but in general not necessary) condition for the squarefreeness of a morphism is known since 1979.

THEOREM 4.1 (Bean et al. [**16**]). *A morphism $\varrho\colon \mathbf{A}^* \to \mathbf{B}^*$ is squarefree if*

(i) *$\varrho(w)$ is squarefree for every squarefree word $w \in \mathbf{A}^*$ of length $|w| \leq 3$;*

(ii) *$a = b$ whenever $a, b \in \mathbf{A}$ and $\varrho(a)$ is a factor of $\varrho(b)$.*

□

If the morphism ϱ is uniform, this condition is in fact also necessary, because in this case $\varrho(a)$ being a factor of $\varrho(b)$ implies that $\varrho(a) = \varrho(b)$. If $a, b \in \mathbf{A}$ exist with $a \neq b$ and $\varrho(a) = \varrho(b)$, then clearly ϱ is not squarefree since $\varrho(ab) = \varrho(a)\varrho(b)$ is a square. This gives the following corollary.

COROLLARY 4.2. *A uniform morphism $\varrho : \mathbf{A}^* \to \mathbf{B}^*$ is squarefree if and only if $\varrho(w)$ is squarefree for every squarefree word $w \in \mathbf{A}^*$ of length $|w| \leq 3$.* □

This corollary corresponds to Brandenburg's Theorem 2 [**18**] which only demands that $\varrho(w)$ is squarefree for every squarefree word $w \in \mathbf{A}^*$ of length exactly 3. A short calculation reveals that this condition is equivalent to (i), because every squarefree word of length smaller than 3 occurs as a factor of a squarefree word of length 3.

For the next characterisation, we need the notion of a pre-square with respect to a morphism ϱ. Let $\mathbf{A}$ be an alphabet, $w \in \mathbf{A}^*$ a squarefree word and $\varrho\colon \mathbf{A}^* \to \mathbf{B}^*$ a morphism. A factor $u \neq \varepsilon$ of $\varrho(w) = \alpha u \beta$ is called a *pre-square* with respect to ϱ, if there exists a word $v \in \mathbf{A}^*$ satisfying: wv is squarefree and u is a prefix of $\beta\varrho(v)$ or vw is squarefree and u is a suffix of $\varrho(v)\alpha$. Obviously, if u is a pre-square, then either $\varrho(wv)$ or $\varrho(vw)$ contains u^2 as a factor.

THEOREM 4.3 (Crochemore [**25**]). *A morphism $\varrho\colon \mathbf{A}^* \to \mathbf{B}^*$ is squarefree*
if and only if

(i) *$\varrho(w)$ is squarefree for every squarefree word $w \in \mathbf{A}^*$ of length $|w| \leq 3$;*
(ii) *for any $a \in \mathbf{A}$, $\varrho(a)$ does not have any internal pre-squares.*

□

It follows that, for a ternary alphabet $\mathbf{A}_3$, a finite test-set exists, as specified in the following corollary. However, the subsequent theorem shows

that, as soon as we consider an alphabet with $|\mathbf{A}| > 3$, no such finite test-sets exist, so the situation becomes more complex when considering larger alphabets.

COROLLARY 4.4 (Crochemore [**25**]). *A morphism $\varrho\colon \mathbf{A}_3^* \to \mathbf{B}^*$ is squarefree if and only if $\varrho(w)$ is squarefree for every squarefree word $w \in \mathbf{A}_3^*$ of length $|w| \leq 5$.* □

THEOREM 4.5 (Crochemore [**25**]). *Let $|\mathbf{A}| > 3$. For any integer n, there exists a morphism $\varrho\colon \mathbf{A}^* \to \mathbf{B}^*$ which is not squarefree, but maps all squarefree words of length up to n on squarefree words.* □

4.2.2. Cubefree and k-powerfree Morphisms. We now move on to characterisations of cubefree and k-powerfree morphisms for $k > 3$. We start with a recent result on binary cubefree morphisms.

THEOREM 4.6 (Richomme, Wlazinski [**76**]). *A set $T \subset \mathbf{A}_2^* = \{0,1\}^*$ is a test-set for cubefree morphisms from $\mathbf{A}_2^*$ to $\mathbf{B}^*$ with $|\mathbf{B}| \geq 2$ if and only if T is cubefree and* $\mathrm{Fact}(T) \supset T_{\min}$*, where*

$$T_{\min} := \{0110110, 1001001, 010110, 101001, 011010, 100101, 00110,$$
$$11001, 01100, 10011, 01010, 10101\}.$$

□

Obviously, the set $T_{\min}$ itself is a test-set for cubefree binary morphisms. Another test-set is the set of cubefree words of length 7, as each word of $T_{\min}$ appears as a factor of this set. There are even single words which contain all the elements of $T_{\min}$ as factors. For instance, the cubefree word

$$0011010110110011001010011$$

is one of the 56 words of length 24 which are test-sets for cubefree morphisms on $\mathbf{A}_2$. The length of this word is optimal: no cube-free word of length 23 contains all the words of $T_{\min}$ as factors.

The following sufficient characterisation of k-powerfree morphisms generalises Theorem 4.1 to integer powers $k > 2$.

THEOREM 4.7 (Bean et al. [**16**]). *Let $\varrho\colon \mathbf{A}^* \to \mathbf{B}^*$ be a morphism for alphabets $\mathbf{A}$ and $\mathbf{B}$ and let $k > 2$. Then ϱ is k-powerfree if*

(i) *$\varrho(w)$ is k-powerfree whenever $w \in \mathbf{A}^*$ is k-powerfree and $|w| \leq k+1$;*

(ii) *$a = b$ whenever $a, b \in \mathbf{A}$ with $\varrho(a)$ a factor of $\varrho(b)$;*

(iii) *the equality $x\varrho(a)y = \varrho(b)\varrho(c)$, with $a, b, c \in \mathbf{A}$ and $x, y \in \mathbf{B}^*$, implies that either $x = \varepsilon$, $a = b$ or $y = \varepsilon$, $a = c$.*

□

As in the squarefree case above, a uniform morphism ϱ for which (i) holds also meets (ii), because uniformity implies that $\varrho(a) = \varrho(b)$. If $a \neq b$, the word $a^{k-1}b$ is k-powerfree but $\varrho(a^{k-1}b) = \varrho(a)^k$ is a k-power, which produces a contradiction. The condition (iii) means that, for all letters $a \in \mathbf{A}$, the images $\varrho(a)$ do not occur as an inner factors of $\varrho(bc)$ for any $b, c \in \mathbf{A}$. In general, this is not a necessary condition for uniform morphisms; an example is given by the Thue-Morse morphism ϱ of (0.1). For instance, $\varrho(00) = 0101 = 0\varrho(1)1$, which violates condition (iii) in Theorem 4.7. Nevertheless, the Thue-Morse morphism is cubefree, even overlapfree, see [**84, 55**].

Alphabets with $|\mathbf{B}| < 2$ only provide trivial results, because the only k-powerfree morphism from $\mathbf{A}^*$ to $\{\varepsilon\}^*$ is the empty morphism ε, and for $|\mathbf{B}| = 1$ the only additional morphism is the map for $|\mathbf{A}| = 1$ that maps the single element in $\mathbf{A}$ to the single letter in $\mathbf{B}$. From now on, we consider alphabets with $|\mathbf{B}| \geq 2$. First, we deal with the case $|\mathbf{A}| \geq 3$.

THEOREM 4.8 (Richomme, Wlazinski [**76**]). *Given two alphabets $\mathbf{A}$ and $\mathbf{B}$ such that $|\mathbf{A}| \geq 3$ and $|\mathbf{B}| \geq 3$, and given any integer $k \geq 3$, there is no finite test-set for k-powerfree morphisms from $\mathbf{A}^*$ to $\mathbf{B}^*$.* □

This again is a negative result, which shows that the general situation is difficult to handle. In general, no finite set of words suffices to verify

the k-powerfreeness of a morphism. The situation improves if we restrict ourselves to uniform morphisms, and look for test-sets for this restricted class of morphisms only.

The existence of finite test-sets of uniform morphisms was recently established by Richomme and Wlazinski [**77**]. Let $|\mathbf{A}| \geq 2$ and $k \geq 3$ be an integer. Define

$$\mathcal{T}^{(k)}(\mathbf{A}) := \mathcal{U}^{(k)}(\mathbf{A}) \cup (\mathcal{F}^{(k)}(\mathbf{A}) \cap \mathcal{V}^{(k)}(\mathbf{A}))$$

where $\mathcal{U}^{(k)}(\mathbf{A})$ is the set of k-powerfree words over $\mathbf{A}$ of length at most $k+1$, and $\mathcal{V}^{(k)}(\mathbf{A})$ is the set of words over $\mathbf{A}$ that can be written in the form

$$a_0 w_1 a_1 w_2 \dots a_{k-1} w_k a_k$$

with letters $a_0, a_1, \dots, a_k \in \mathbf{A}$ and words $w_1, w_2 \dots w_k \in \mathbf{A}^*$, which contain every letter of $\mathbf{A}$ at most once and satisfy $\big||w_i| - |w_j|\big| \leq 1$. Obviously, this set is finite and comprises words with a maximum length of

$$\max\{|w| \mid w \in \mathcal{T}^{(k)}(\mathbf{A})\} \leq k(|\mathbf{A}|) + 1) + 1.$$

THEOREM 4.9 (Richomme, Wlazinski [**77**]). *Let $|\mathbf{A}| \geq 2$ and $k \geq 3$ be an integer. The finite set $\mathcal{T}^{(k)}(\mathbf{A})$ is a test-set for k-powerfreeness of uniform morphisms on $\mathbf{A}^*$.* □

Due to the upper bound on the maximum length of words in $\mathcal{T}^{(k)}(\mathbf{A})$, the following corollary is immediate.

COROLLARY 4.10 (Richomme, Wlazinski [**77**]). *A uniform morphism ϱ on $\mathbf{A}^*$ is k-powerfree for an integer power $k \geq 3$ if and only if $\varrho(w)$ is k-powerfree for all k-powerfree words w of length at most $k(|\mathbf{A}|+1)+1$.* □

Although this result provides an explicit test-set for k-powerfreeness, it is of limited practical use, simply because the test-set becomes large very quickly. Already for $\mathbf{A} = \mathbf{A}_4$ and $k = 3$, the set $\mathcal{T}^{(3)}(\mathbf{A}_4)$ has 26247020 elements. For comparison, the set of cubefree words in four letters of length

16, as required in Corollary 4.10, has 1939267560 elements, so is still much larger.

Finally, let us quote the following result of Keränen [**45**], which characterises k-powerfree binary morphisms and indicates that the test-set of Theorem 4.9 is far from optimal. Note that a word $w \in \mathbf{A}^*$ is called *primitive*, if $w = v^k$, with $v \in \mathbf{A}^*$ and $k \in \mathbb{N}$, implies that $k = 1$, meaning that w is not a proper power of another word v.

THEOREM 4.11 (Keränen [**45**]). *Let $\varrho\colon \mathbf{A}_2^* \to \mathbf{B}^*$ be a uniform morphism with $\varrho(0) \neq \varrho(1)$ and primitive words $\varrho(0), \varrho(1)$ and $\varrho(01)$. For every word $w \in \mathcal{F}^{(k)}(\mathbf{A}_2)$, $\varrho(w)$ is k-powerfree if and only if $\varrho(v)$ is k-powerfree for every subword v of w with*

$$|v| \leq \begin{cases} 4 & \text{for } 3 \leq k \leq 6; \\ \frac{2}{3}(k+1) & \text{for } k \geq 7. \end{cases}$$

4.3. Combinatorial Entropy

Let $\mathbf{A}$ be an alphabet. A subset $S \subset \mathbf{A}^*$ is called *factorial* if for any word $w \in S$ all factors of w are also contained in S. Define for a factorial set $S \subset \mathbf{A}^*$ the number of words of length n occurring in S by $c_S(n) := |S(n)|$. This number gives some idea of the complexity of S: the larger the number of words of length n, the more diverse and complicated the set. That is why $c_S : \mathbb{N} \to \mathbb{N}$ is called the *complexity function* of S. As S is factorial we infer that

$$c_S(m+n) \leq c_S(m)c_S(n). \tag{4.7}$$

DEFINITION 4.12. *The* combinatorial entropy *of an infinite factorial set $S \subset \mathbf{A}^*$ is defined by*

$$h(S) = \lim_{n\to\infty} \tfrac{1}{n} \log c_S(n).$$

The inequality (4.7) ensures that this limit exists, see for example [**4**, Lemma 1]. This is also known as Fekete's Lemma, see [**81**, Lemma 1.2.1], which says in addition that $h(S) = \inf_{n\in\mathbb{N}} \frac{1}{n} \log c_S(n)$. We note the following:

(i) If $S \subset \mathbf{A}^*$ with $|\mathbf{A}| = \ell$, then $1 \leq c_S(n) \leq \ell^n$ for all n which implies $0 \leq h(S) \leq \log(\ell)$.

(ii) If $S = \mathbf{A}^*$ with $|\mathbf{A}| = \ell$, then $c_S(n) = \ell^n$ and $h(S) = \log(\ell)$.

4.3.1. Explicit Cases. Explicitly, we intend to derive upper and lower bounds for the entropies of the following sets, which are obviously factorial. The numerical results are presented in the corresponding sections of Chapter 7.

7.1 $\mathcal{F}^{(3)}(\mathbf{A}_2)$ - Cubefree words over $\mathbf{A}_2 = \{0, 1\}$

7.2 $\mathcal{F}^{(2)}(\mathbf{A}_3)$ - Squarefree words over $\mathbf{A}_3 = \{0, 1, 2\}$

7.3 $\mathcal{F}^{(>\frac{7}{4})}(\mathbf{A}_3)$ - $(\frac{7}{4})^+$-powerfree words over $\mathbf{A}_3$

7.4 $\mathcal{F}^{(>\frac{7}{3})}(\mathbf{A}_2)$ - $(\frac{7}{3})^+$-powerfree words over $\mathbf{A}_2$

7.5 $\mathcal{F}^{(2)}(\mathbf{A}_4)$ - Squarefree words over $\mathbf{A}_4 = \{0, 1, 2, 3\}$

7.6 $\mathcal{F}^{(>\frac{7}{5})}(\mathbf{A}_4)$ - $(\frac{7}{5})^+$-powerfree words over $\mathbf{A}_4$

As already mentioned the Cases 7.1 and 7.2 are the two classical cases. The Case 7.5 is included to complete the picture.

Dejean [**30**] and later Brandenburg [**18**] introduced a *repetition threshold* $\mathrm{RT}(\ell)$, defined as the smallest number $\alpha \in \mathbb{R}$ such that there exists an infinite word over $\mathbf{A}_\ell$ that is α^+-powerfree. Dejean proved that every sufficiently long word over $\mathbf{A}_3$ contains a $\frac{7}{4}$-power, and that there are infinite words over $\mathbf{A}_3$ which are $\frac{7}{4}^+$-powerfree. This implies that $\mathrm{RT}(3) = \frac{7}{4}$. Moreover, she conjectured that

$$\mathrm{RT}(\ell) = \begin{cases} \frac{7}{4}, & \ell = 3, \\ \frac{7}{5}, & \ell = 4, \\ \frac{\ell}{\ell-1}, & \text{otherwise.} \end{cases}$$

This conjecture was proved for $\ell = 2$ by Thue [**84, 55**], for $\ell = 4$ by Pansiot [**69**], for $5 \leq \ell \leq 11$ by Ollagnier [**68**], for $12 \leq \ell \leq 14$ by Mohammad-Noori and Currie [**58**], for $\ell \geq 33$ by Carpi [**20**], for $\ell \geq 27$ by Currie and Rampersad [**29, 27**] and for the remaining cases independently by Currie and Rampersad [**28**] as well as Rao [**73**].

Dejean's conjecture implies that every infinite or sufficiently long word over $\mathbf{A}_\ell$ contains an α-power with $\alpha \geq \mathrm{RT}(\ell)$. As $\mathrm{RT}(\ell)$-powers are unavoidable we call words over $\mathbf{A}_\ell$ which are $\mathrm{RT}(\ell)^+$-powerfree *minimally repetitive*. Obviously, the set of all minimally repetitive words is of particular interest. Ochem [**64**] proved the exponential growth of the number of elements of the set of minimally repetitive words over $\mathbf{A}_3$ and $\mathbf{A}_4$. This implies that the corresponding entropies are non-zero and in Case 7.3 and Case 7.6 we try to compute their lower bounds by means of Kolpakov's method [**49**].

Karhumäki and Shallit showed in [**43**] that for $\mathbf{A}_2$ the dividing line between polynomial and exponential growth is $\frac{7}{3}$. This means that the set of binary minimally repetitive, i.e. 2^+-powerfree, words grows only polynomially and hence has zero entropy. Therefore, we deal in Case 7.4 with the set of binary words which are $\frac{7}{3}^+$-powerfree. Since $\frac{7}{3}$-powers are unavoidable for exponential growth, a $\frac{7}{3}^+$-powerfree word is also referred to as *binary quasi-minimally repetitive* word.

4.3.2. Lower Bounds for the Entropy via Powerfree Morphisms. For the rest of this chapter, if not specified otherwise, let $\mathcal{F}$ stand for an infinite set of k-powerfree or α^+-powerfree words. A word $w \in \mathcal{F}$ is called *powerfree* in both cases and it will be clear from the context what is meant. Since the set $\mathcal{F}$ is obviously factorial, the entropy $h(\mathcal{F})$ exists. Until very recently, all methods used to prove that $h(\mathcal{F})$ is positive and to establish lower bounds were based on powerfree morphisms. Clearly, a powerfree morphism, iterated on a single letter, produces powerfree words of increasing length and suffices to show the existence of infinite powerfree words. For example, the fact that the Thue-Morse morphism (0.1) is overlapfree and hence cubefree

shows the existence of overlapfree and cubefree words of arbitrary length in two letters.

To prove that the entropy is actually positive, one has to show that the number of powerfree words grows exponentially with their length. Essentially, this is achieved by considering powerfree morphisms from a larger alphabet. The following theorem is a generalisation of Brandenburg's method, compare [**18**], and provides a path to produce lower bounds for the entropy of k-powerfree words.

THEOREM 4.13. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be alphabets with* $|\mathbf{A}| = t|\mathbf{B}|$*, where* $t > 1$ *is an integer. If there exists an* r*-uniform* k*-powerfree morphism* $\varrho\colon \mathbf{A}^* \to \mathbf{B}^*$*, then*

$$h(\mathcal{F}^{(k)}(\mathbf{B})) \geq \frac{\log t}{r-1}.$$

PROOF. For this proof define $h := h(\mathcal{F}^{(k)}(\mathbf{B}))$, $c(n) := c_{\mathcal{F}^{(k)}(\mathbf{B})}(n)$ and $s := |\mathbf{B}|$. Label the elements of $\mathbf{A}$ as $\{a_{11}, \ldots, a_{1t}, a_{21}, \ldots, a_{2t}, \ldots, a_{s1}, \ldots, a_{st}\}$ and the elements of $\mathbf{B}$ as $\{b_1, \ldots, b_s\}$. Define a map $\varphi : \mathbf{A}^* \to \mathbf{B}^*$ as $\varphi(a_{ij}) := b_i$ for $i = 1, \ldots, s$ and $j = 1, \ldots, t$. Hence $|\varphi^{-1}(b_i)| = t$. Every k-powerfree word of length m over $\mathbf{B}$ has t^m different preimages under φ which, by construction, consist only of k-powerfree words. These words are mapped by ϱ, which is injective due to its k-powerfreeness, to different k-powerfree words of length mr over $\mathbf{B}$. This implies the inequality

$$c(mr) \geq t^m c(m) \tag{4.8}$$

for any $m > 0$ and means that

$$\left(\frac{c(mr)}{c(m)}\right)^{\frac{1}{m}} \geq t.$$

Hence

$$r\,\frac{\log c(mr)}{mr} - \frac{\log c(m)}{m} \geq \log t$$

for any $m > 0$. Taking the limit as $m \to \infty$ gives

$$(r-1)h \geq \log t,$$

thus establishing the lower bound. □

This result means that, whenever we find a uniform k-powerfree morphism from a sufficiently large alphabet, we have found a lower bound for the entropy, and in particular we have shown that the entropy is strictly positive. Clearly, the larger t and the smaller r the better the bound, so one is particularly interested in uniform k-powerfree morphisms from large alphabets of minimal length.

Another method due to Brinkhuis [**19**], which is related to Brandenburg's method, can be generalised as follows. Let again $\mathbf{B} = \{b_1, \ldots, b_s\}$ be an alphabet and $t \in \mathbb{N}$. For $i \in \{1, \ldots s\}$ let

$$\mathscr{U}_i := \{U_{i,1}, U_{i,2}, \ldots, U_{i,t}\}$$

with $U_{i,j} \subset \mathcal{F}^{(k)}(\mathbf{B}) \subset \mathbf{B}^*(r)$. The set $\mathscr{U} = \{\mathscr{U}_1, \ldots \mathscr{U}_n\}$ is called an (k, r, t)-*Brinkhuis-set* if the r-uniform substitution (in the context of formal language theory), compare for example [**17**, Sec. 3.2], ϱ from $\mathbf{B}^*$ to itself defined by

$$\varrho\colon b_i \mapsto \mathscr{U}_i \quad \text{for} \quad i = 1, \ldots, s$$

has the property $\varrho(\mathcal{F}^{(k)}(\mathbf{B})) \subset \mathcal{F}^{(k)}(\mathbf{B})$. In other words $\mathscr{U}$ is an (k, r, t)-Brinkhuis-set if the substitution of every letter b_i, occurring in a k-powerfree word, by an element of $\mathscr{U}_i$ results in a k-powerfree word over $\mathbf{B}$. The existence of a (k, r, t)-Brinkhuis-set delivers the lower bound

$$h(\mathcal{F}^{(k)}(\mathbf{B})) \geq \frac{\log t}{r-1}$$

because every k-powerfree word of length m is mapped to t^m powerfree words of length rm; compare (4.8).

The method of Brinkhuis is stronger than the method of Brandenburg. Not every (k, r, t)-Brinkhuis-set implies a map according to Theorem 4.13, since we only choose one word out of $\mathcal{U}_i$, see [**17**, p. 287] for an example. Conversely, if there exists an r-uniform k-powerfree morphism $\varrho\colon \mathbf{A}^* \to \mathbf{B}^*$ according to Theorem 4.13, then there exists a (k, r, t)-Brinkhuis-set, namely $\mathscr{U}_i = \{\varrho(a_{i1}), \ldots, \varrho(a_{it})\}$ for $i = 1, \ldots, s$, with the notation of Theorem 4.13.

Brinkhuis' method was applied in [**36, 82**]; see also Sections 7.1 and 7.2 for a summary of bounds obtained for binary cubefree and ternary squarefree words. These bounds have in common that they are nowhere near the actual value of the entropy, and while a systematic improvement is possible by increasing the value of t in Theorem 4.13 (which, however, also means that one has to consider larger values of r), it will always result in a much smaller growth rate, because only a subset of words is obtained in this way.

In 2007 Kolpakov introduced a completely different approach [**49**], which we will explain and generalise in Chapter 6. The lower bounds derived by this approach are much better, in fact they are the best known so far.

4.3.3. Upper Bounds for the Entropy. A simple way to provide upper bounds on the entropy of an infinite factorial set $S \subset \mathbf{A}_\ell^*$ is based on the enumeration of the elements of S up to some length. Clearly, the number of words $c(n) := c_S(n)$ is bounded by ℓ^n, so the corresponding entropy is $h(S) \le \log \ell$, as mentioned above. Suppose we know the actual value of $c(n)$ for some fixed n. Then, due to the factorial nature of the set S,

$$c(mn) \le c(n)^m$$

for any $m \ge 1$. Hence

$$h(S) = \lim_{m\to\infty} \frac{\log c(mn)}{mn} \le \frac{\log c(n)}{n}, \tag{4.9}$$

which, for any n, yields an upper bound for h. Obviously, the larger the value of n, the better the bound obtained in this way. In some cases, the

bound can be slightly improved by considering words that overlap in a couple of letters; see [**4**] for an example.

Recall that $\mathcal{F}$ stands for an infinite set of k-powerfree or α^+-powerfree words. In this case (4.9) leads to the following approximation of the entropy. The sets

$$\mathcal{S}_m := \mathcal{S}_m(\mathbf{A}) = \{w \in \mathbf{A}^* \mid \text{ every factor of } w \text{ with length } \leq m \text{ is powerfree}\}$$

form an *outer approximation* of $\mathcal{F}$, i.e. $\mathcal{S}_n \subset \mathcal{S}_m$ for every pair $m, n \in \mathbb{N}$ with $n \geq m$ and

$$\bigcap_{m=1}^{\infty} \mathcal{S}_m = \mathcal{F}. \tag{4.10}$$

Note that obviously the sets $\mathcal{S}_m$ are factorial and $\mathcal{F}(n) = \mathcal{S}_m(n)$ for all $n \leq m$.

LEMMA 4.14. *Let $\mathcal{F}$ stand for an infinite set of k-powerfree or α^+-powerfree words and let $\mathcal{S}_m$ be an outer approximation of $\mathcal{F}$. Then*

$$\lim_{m\to\infty} h(\mathcal{S}_m) = h(\mathcal{F}).$$

PROOF. We will show that

$$h(\mathcal{F}) \leq h(\mathcal{S}_m) \leq h(\mathcal{F}) + \varepsilon$$

for any $\varepsilon > 0$ and all $m > M \in \mathbb{N}$. Since $\mathcal{F} \subset \mathcal{S}_m$ we know that $h(\mathcal{F}) \leq h(\mathcal{S}_m)$ for all $m \in \mathbb{N}$. By definition

$$\lim_{n\to\infty} \tfrac{1}{n} \log(|\mathcal{F}(n)|) = h(\mathcal{F}),$$

so there is an $M \geq 1$ such that $\frac{1}{M} \log(|\mathcal{F}(M)|) < h(\mathcal{F}) + \varepsilon$. For all $m \geq M$ we know that $\mathcal{S}_m(M) = \mathcal{F}(M)$ and hence with (4.9)

$$h(\mathcal{S}_m) = \lim_{n\to\infty} \tfrac{1}{n} \log(|\mathcal{S}_m(n)|) \leq \tfrac{1}{M} \log|\mathcal{S}_m(M)| = \tfrac{1}{M} \log|\mathcal{F}(M)| < h(\mathcal{F}) + \varepsilon.$$

□

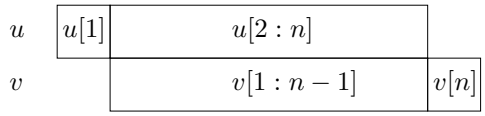


FIGURE 4.1. v as a *descendant* of u or u as an *ancestor* of v.

There are several methods to calculate sharper upper bounds for ternary squarefree and binary cubefree words, see for example [**63, 75, 67, 37**], which could be generalised. However, we prefer to present here a general method to derive upper bounds for the entropy which, at the same time, provides the first steps for the derivation of the lower bound in Chapter 6. We will see in Chapter 7 that, compared to other methods, it delivers upper bounds which are only slightly greater. The following definitions will be central.

DEFINITION 4.15. *Let $n > 1$ and $u, v \in \mathcal{F}(n)$. We call v a* descendant *of u and u an* ancestor *of v if*

$$u[2:n] = v[1:n-1] \quad \textit{and} \quad uv[n] = u[1]v \in \mathcal{F}(n+1),$$

see Figure 4.1 for an illustration. Moreover, we call v a quasi-descendant *of u and u a* quasi-ancestor *of v, if v is isomorphic to some descendant of u, i.e. there is a permutation of letters $\varrho : \mathbf{A}^* \to \mathbf{A}^*$, such that $\varrho(v)$ is a descendant of u or equivalently u is an ancestor of $\varrho(v)$.*

DEFINITION 4.16. *A word $w \in \mathcal{F}(n)$ is called* open *if the following holds:*

(1) *w has at least one ancestor and one descendant.*
(2) *At least one ancestor and one descendant of w is open.*

For example the word $020 \in \mathcal{F}^{(2)}(\mathbf{A}_3)$ is open, since we have the following line of descendants: $020, 201, 010, 102, 020$.

A word which is not open, is called *closed* in the following sense.

DEFINITION 4.17. *A word $w \in \mathcal{F}(n)$ is called* right closed (left closed)*, if w has no descendant (ancestor) or all descendants (ancestors) of w are*

right closed (left closed). A word w is called closed *if it is right closed or left closed.*

For example the word $0102010 \in \mathcal{F}^{(2)}(\mathbf{A}_3)$ is right and left closed.

Let $3 \leq m \in \mathbb{N}$ and recall from (4.6) that $\mathcal{F}'(m)$ stands for the set of equivalence classes of powerfree words of length m. Now, we define the set of open words among them as

$$\mathcal{F}''(m) := \{w \in \mathcal{F}'(m) \mid w \text{ is open } \} = \{w_1, \ldots, w_s\}, \tag{4.11}$$

where we think of this set as ordered lexicographically.

Definition 4.18. *Let $3 \leq m \in \mathbb{N}$. For $\mathcal{F}''(m) = \{w_1, \ldots, w_s\}$ from (4.11) define an $s \times s$-matrix $\Delta_m = (\delta_{ij})$, where*

$$\delta_{ij} := \begin{cases} 1, & \text{if } w_i \text{ is a quasi-ancestor of } w_j \\ 0, & \text{otherwise} \end{cases}.$$

As Δ_m is a non-negative matrix it possesses a largest eigenvalue $\lambda_m \geq 0$, which is called Perron eigenvalue*, see for example* [**54**, Lemma 4.4.3].

A non-negative square matrix M is called *irreducible*, if for each ordered pair of indices (i, j), there exists some $n = n(i, j) > 0$ such that $(M^n)_{i,j} > 0$. In this thesis we adopt the convention that for any matrix $M^0 = id$ and so the 1×1 matrix (0) is irreducible. Furthermore, a non-negative square matrix M is called *primitive*, if $M^n > 0$ for some $n > 0$, i.e. every entry of M^n is strictly greater than 0. For example by [**54**, Definition 4.5.7, Theorem 4.5.8] we know that a matrix M is primitive if and only if M is irreducible and *aperiodic* , i.e. for an index i

$$\gcd\{n \mid (M^n)_{ii} > 0\} = 1.$$

We will show in Chapter 7 that for the examples we look at, the matrix Δ_m is primitive since it is irreducible and aperiodic. However, we will see in Chatper 6 that we actually only need that Δ_m is irreducible. Already

the Perron-Frobenius Theorem for irreducible matrices, see for example [**54**, Theorem 4.2.3] or [**46**, Theorem 1.3.5], tells us what we need. Namely, the largest eigenvalue λ_m of Δ_m, the Perron eigenvalue, is geometrically as well as algebraically simple and possesses a strictly positive right eigenvector, which is unique up to multiplication with a positive factor. Moreover, the Perron eigenvalue $\lambda_m > 0$ is then the only eigenvalue of Δ_m with a non-negative eigenvector.

We conclude this chapter with the following theorem, which shows how all these definitions lead to an upper bound for the entropy of powerfree words. The theorem will be proved in the course of Chapter 5, as it will be easier to prove it in the context of symbolic dynamics.

THEOREM 4.19. *Let $\mathcal{F}$ be an infinite set of either k-powerfree or α^+-powerfree words. For $3 \leq m \in \mathbb{N}$, let Δ_m be the matrix from Definition 4.18. If we denote the Perron eigenvalue of Δ_m by λ_m, then*

$$\log(\lambda_m) \geq h(\mathcal{F}).$$

□

CHAPTER 5

Dynamical Aspects

In this chapter powerfree words are described by means of symbolic dynamics. It is shown that the entropy of a powerfree shift space coincides with the combinatorial entropy of the corresponding set of powerfree words. Moreover, we define dynamical systems which arise from substitutions. Although these systems are suitable to show that there are infinitely many powerfree words, we show that their entropy is zero. We conclude with a section on topological entropy and show explicitly that for shift spaces the combinatorial and topological entropy coincide.

5.1. Shift Spaces

5.1.1. Definitions. A *bi-infinite* sequence of letters from an alphabet $\mathbf{A}$ is a map

$$x : \mathbb{Z} \to \mathbf{A},\ i \mapsto x_i$$

and is denoted by $x = (x_i)_{i\in\mathbb{Z}}$ or $x = \ldots x_{-2}x_{-1}x_0x_1x_2\ldots$, compare [**54**]. The symbol x_i is called the ith *coordinate* of x. We write a specific sequence with a dot between the -1st and the 0th coordinate. For example

$$x = \ldots 01.32\ldots$$

means that $x_{-2} = 0, x_{-1} = 1, x_0 = 3, x_1 = 2$ and so on. The set of all bi-infinite sequences of letters from $\mathbf{A}$ is called the *full* $\mathbf{A}$*-shift* and is denoted by

$$\mathbf{A}^{\mathbb{Z}} = \{x = (x_i)_{i\in\mathbb{Z}} \mid x_i \in \mathbf{A} \text{ for all } i \in \mathbb{Z}\}.$$

Each $x \in \mathbf{A}^{\mathbb{Z}}$ is called a *point* of the full $\mathbf{A}$-shift. The *shift map* $\sigma : \mathbf{A}^{\mathbb{Z}} \to \mathbf{A}^{\mathbb{Z}}$ maps a point x to the point $y = \sigma(x)$ whose ith coordinate is $y_i = x_{i+1}$.

Thus σ shifts every coordinate one place to the left. Of course there is also the inverse operation σ^{-1} of shifting one place to the right.

If $x \in \mathbf{A}^{\mathbb{Z}}$ and w is a finite word over $\mathbf{A}$, it is said that w *occurs in* or *is contained in* x if there are indices i and j such that $w = x[i : j]$. Note that the empty word ε occurs in every point x, since $\varepsilon = x[1 : 0]$. Let F be a set of finite words over $\mathbf{A}$, which we will consider as the set of *forbidden words*. For any such F, define X_F to be the subset of elements of $\mathbf{A}^{\mathbb{Z}}$ which do not contain any word of F. A *shift space* is a subset X of the full shift $\mathbf{A}^{\mathbb{Z}}$ of the form $X = X_F$ for some set F of forbidden words over $\mathbf{A}$. For a given shift space X_F there may be several different sets describing it. Note that the empty set $\varnothing$ is a shift space, since F can be $\mathbf{A}^*$. We say that a shift space X is of *finite type* if there is a finite set F such that $X = X_F$. In any case, even if F is infinite, it is countable since its elements can be arranged in a list.

Clearly, shift spaces are *shift invariant*, meaning that $\sigma(X_F) = X_F$. When a shift space X is a subset of a shift space Y, it is called a *subshift* of Y. We denote the set of words or blocks of length n which occur in some point of a shift space $X \subset \mathbf{A}^{\mathbb{Z}}$ as

$$\mathcal{B}_n(X) := \{w \in \mathbf{A}^* \mid |w| = n, w \text{ occurs in some } x \in X\}.$$

The set of all words occurring in some point of a shift space X is denoted by

$$\mathcal{B}(X) := \bigcup_{n=1}^{\infty} \mathcal{B}_n(X)$$

and is also called the *language* of X. Obviously, $\mathcal{B}(X)$ is factorial and hence its entropy, compare Definition 4.12, exists.

5.1.2. Powerfree Shift Spaces. For $k \in \mathbb{N}$ and $\alpha \in \mathbb{Q}$ we define the set of words over $\mathbf{A}$ which contain at least one k-power respectively one α^+-power as a factor as

$$\mathcal{C}^{(k)}(\mathbf{A}) := \mathbf{A}^* \setminus \mathcal{F}^{(k)}(\mathbf{A}) \quad \text{and} \quad \mathcal{C}^{(>\alpha)}(\mathbf{A}) := \mathbf{A}^* \setminus \mathcal{F}^{(>\alpha)}(\mathbf{A}). \tag{5.1}$$

Their corresponding shift spaces $X_{\mathcal{C}^{(k)}(\mathbf{A})}$ and $X_{\mathcal{C}^{(>\alpha)}(\mathbf{A})}$ consists of all k-powerfree respectively α^+-powerfree elements of $\mathbf{A}^{\mathbb{Z}}$. Note that these shift spaces are of infinite type. To simplify the notation we write from now on just $\mathcal{F}$, if we mean the sets $\mathcal{F}^{(k)}(\mathbf{A})$ and $\mathcal{F}^{(>\alpha)}(\mathbf{A})$. Analogously, $\mathcal{C}$ stands for $\mathcal{C}^{(k)}(\mathbf{A})$ and $\mathcal{C}^{(>\alpha)}(\mathbf{A})$ simultaneously.

Although the shift spaces $X_{\mathcal{C}}$ are not of finite type, we can approximate them by shifts of finite type. Observe that $\mathcal{C}(m)$ denotes the set of words of length m which contain at least one forbidden power as a factor and that this set is finite. The shift space $X_{\mathcal{C}(m)}$ consists of all points from $\mathbf{A}^{\mathbb{Z}}$ whose factors of length m are powerfree. Obviously, for $m \geq n$, we have $X_{\mathcal{C}(m)} \subset X_{\mathcal{C}(n)}$ and

$$X_{\mathcal{C}} = \bigcap_{m=1}^{\infty} X_{\mathcal{C}(m)} = X_{\cup_{m=1}^{\infty}\mathcal{C}(m)}. \tag{5.2}$$

Note that not necessarily all elements of $\mathcal{F}$ occur in points of the corresponding shift spaces. We can characterise them as follows.

LEMMA 5.1. *Let $\mathcal{F}$ and $\mathcal{C}$ stand for $\mathcal{F}^{(k)}(\mathbf{A})$ and $\mathcal{C}^{(k)}(\mathbf{A})$ or $\mathcal{F}^{(>\alpha)}(\mathbf{A})$ and $\mathcal{C}^{(>\alpha)}(\mathbf{A})$. A word $w \in \mathcal{F}(m)$ is open if and only if $w \in \mathcal{B}_m(X_{\mathcal{C}(m+1)})$.* □

For example, the word $0102010 \in \mathcal{F}^{(2)}(\mathbf{A}_3)$, but it cannot occur as a factor of a point in $X_{\mathcal{C}^{(2)}(m)}$ for $m \geq 8$ since it is closed, namely right and left closed at the same time. Moreover, there are words which are only closed to one side. For example, the word $01021201021 \in \mathcal{F}^{(2)}(\mathbf{A}_3)$ is left closed but not right closed, hence it cannot occur as factor of a point in $X_{\mathcal{C}^{(2)}(m)}$ for $m \geq 12$.

Note that the shift space $X_{\mathcal{C}(m+1)}$ is an m-step shift of finite type, i.e. its set of forbidden words consists of words of length $m+1$ only. $X_{\mathcal{C}(m+1)}$ can be described by the following graph, compare [**54**, Theorem 2.3.2.]: The vertices are the elements of $\mathcal{B}_m(X_{\mathcal{C}(m+1)}) = \{w_1, \ldots, w_r\}$. Two vertices u, v are linked by an edge, if u is an ancestor of v. Let $\Gamma_m = (\gamma_{ij})$ be the

adjacency matrix of this graph, i.e.

$$\gamma_{ij} := \begin{cases} 1, & \text{if } w_i \text{ is a ancestor of } w_j \\ 0, & \text{otherwise} \end{cases},$$

and let Γ_m^0 stand for the identity matrix. Note that for $n \geq m$

$$(\Gamma_m^{(n-m)})_{i,j}$$

gives the number of words of length n, occurring in points of $X_{\mathcal{C}(m+1)}$, which have w_i as a prefix and w_j as a suffix. In other words: $(\Gamma_m^{(n-m)})_{i,j}$ counts the number of paths, as a line of descendants, starting with w_i and ending in w_j. Hence we have for $n \geq m$ that

$$|\mathcal{B}_n(X_{\mathcal{C}(m+1)})| = \sum_{i,j=1}^{r} (\Gamma_m^{(n-m)})_{i,j}.$$

Recall the definition of the set $\mathcal{F}''(m) = \{w_1, \ldots, w_s\}$ from (4.11) and note that

$$(\mathcal{B}_m(X_{\mathcal{C}(m+1)}))' = \{w_1', \ldots, w_r'\} = \mathcal{F}''(m).$$

If $e(i)$ for $i \in \{1, \ldots, s\}$ denotes the number of words which are represented by $w_i \in \mathcal{F}''(m)$, we have $\sum_{i=1}^{s} e(i) = r$. Moreover, define

$$e_{max} := \max_{1 \leq i \leq s} e(i) \quad \text{and} \quad e_{min} := \min_{1 \leq i \leq s} e(i).$$

Explicitly, for every word $w \in \mathcal{B}_n(X_{\mathcal{C}(m+1)})$, there are $e(i)$ words in $\mathcal{B}_n(X_{\mathcal{C}(m+1)})$ which are the image of w under a permutation of letters. Remember the matrix Δ_m from Definition 4.18 and note that in this matrix equivalent words are summarised so that we have

$$e_{min} \sum_{i,j=1}^{s} (\Delta_m^{(n-m)})_{i,j} \leq |\mathcal{B}_n(X_{\mathcal{C}(m+1)})| = \sum_{i,j=1}^{r} (\Gamma_m^{(n-m)})_{i,j} \leq e_{max} \sum_{i,j=1}^{s} (\Delta_m^{(n-m)})_{i,j}.$$

We conclude with [**54**, Theorem 4.4.4] that

$$h(X_{\mathcal{C}(m+1)}) \leq \lim_{n\to\infty} \frac{1}{n} \log e_{max} + \lim_{n\to\infty} \frac{1}{n} \log(\sum_{i,j=1}^{s} (\Delta_m^{(n-m)})_{i,j}) = \log \lambda_m$$

and similarly $\log \lambda_m \leq h(X_{\mathcal{C}(m+1)})$, such that

$$h(X_{\mathcal{C}(m+1)}) = \log \lambda_m. \tag{5.3}$$

Now, we are finally in a position to prove Theorem 4.19 on page 116.

PROOF. Recall from (4.10) the outer approximation $\mathcal{S}_m$ of $\mathcal{F}$. First of all we show that

$$h(\mathcal{S}_{m+1}) = h(X_{\mathcal{C}(m+1)}). \tag{5.4}$$

From what we have said so far it is clear that $\mathcal{B}_n(X_{\mathcal{C}(m+1)}) \subset \mathcal{S}_{m+1}(n)$. By Lemma 5.1 the set $\mathcal{B}_n(X_{\mathcal{C}(m+1)})$ only contains open words, in the sense that every factor of length m of an element of $\mathcal{B}_n(X_{\mathcal{C}(m+1)})$ is open. We think of $w \in \mathcal{B}_n(X_{\mathcal{C}(m+1)})$ as $w[1:m]$ followed by the last letter of the $n-m$ descendants of $w[1:m]$.

Let $w \in \mathcal{S}_{m+1}(n) \setminus \mathcal{B}_n(X_{\mathcal{C}(m+1)})$. If $w[1:m]$ is right and left closed, then $w[1:m]$ possess only a finite number of ancestors and descendants. Hence there is an index $N(w[1:m])$ such that w does not occur as a factor in any word of $\mathcal{S}_{m+1}(n)$ for $n > N(w[1:m])$. If $w[1:m]$ is left closed but not right closed, then $w[1:m]$ possess an infinite row of descendants, but only a finite number of ancestors. Let $a(w[1:m])$ be the number of ancestors and $d(w[1:m])$ be the number of the first descendant which is open. This word exists since there is only a finite number of words in $\mathcal{S}_{m+1}(m) = \mathcal{F}(m)$ and with any repetition in the row of descendants we have found an open word. So there are $a(w[1:m]) + m + d(w[1:m]) - 1$ starting positions of words of length n which might not occur in $\mathcal{B}_n(X_{\mathcal{C}(m+1)})$. Analogously, if $w[1:m]$ is right closed but not left closed , let $a(w[1:m])$ be the number of the first ancestor which is open and let $d(w[1:m])$ be the finite number of descendants. With $C := \{w \in \mathcal{F}(m) \mid w \text{ is either left or right closed but not both}\}$ and

$D := \{w \in \mathcal{F}(m) \mid w \text{ is left and right closed}\}$ we have

$$c = \sum_{w \in C} a(w[1:m]) + m + d(w[1:m]) - 1$$

and conclude that $|\mathcal{S}_{m+1}(n) \setminus \mathcal{B}_n(X_{\mathcal{C}(m+1)})| \leq c$ for all $n > N := \max_{w \in D} N(w)$. Hence

$$h(\mathcal{S}_{m+1}) = \lim_{n\to\infty} \tfrac{1}{n} \log |\mathcal{S}_{m+1}(n)| = \lim_{n\to\infty} \tfrac{1}{n} \log(|\mathcal{B}_n(X_{\mathcal{C}(m+1)})| + c) = h(X_{\mathcal{C}(m+1)}).$$

With (4.10) and (5.3) this implies that $\log \lambda_m = h(X_{\mathcal{C}(m+1)}) = h(\mathcal{S}_{m+1}) \geq h(X_{\mathcal{F}})$ which completes the proof. □

By [**54**, Proposition 4.4.6] we know that

$$\lim_{m\to\infty} h(X_{\mathcal{C}(m+1)}) = h(X_{\mathcal{C}}).$$

So we obtain, with Lemma 4.14, the following result.

Corollary 5.2. *Let $\mathcal{F}$ be an infinite set of either k-powerfree or α^+-powerfree words and $\mathcal{C}$ the corresponding set of power containing words. Then*

$$h(\mathcal{F}) = \lim_{m\to\infty} h(\mathcal{S}_m) = \lim_{m\to\infty} h(X_{\mathcal{C}(m+1)}) = h(X_{\mathcal{C}}).$$

□

5.1.3. Shifts as Dynamical Systems. In general, dynamical systems are defined as follows.

Definition 5.3. *A* (topological) dynamical system (M, φ) *consists of a compact topological Hausdorff space M together with a continuous map $\varphi : M \longrightarrow M$. If φ is a homeomorphism we call (M, φ) an* invertible dynamical system.

Let X be a shift space, $u \in \mathcal{B}(X)$ and $i \in \mathbb{Z}$. Define a *cylinder set* over X as

$$C_i^X(u) := \left\{x \in X \mid x_{[i,i+|u|-1]} = u\right\},$$

i.e. $C_i^X(u)$ is the set of points in which the word u occurs starting at position i. The cylinder sets form a countable basis for a topology on the shift space X, see [**46**, Section 1.1]. The following metric generates the topology. For $x, y \in X$ define

$$d(x,y) := \begin{cases} 2^{-i} & \text{if } x \neq y \text{ and } i \text{ is maximal so that } x_{[-i,i]} = y_{[-i,i]} \\ 0 & \text{if } x = y \end{cases}. \tag{5.5}$$

Observe that for any $x \in X$ and any integer $n \geq 0$ we have that

$$\begin{aligned} B_{2^{-(n-1)}}(x) &= \left\{ y \in X \mid d(x,y) < 2^{-(n-1)} \right\} \\ &= \left\{ y \in X \mid y_{[-n,n]} = x_{[-n,n]} \right\} = C_{-n}^X(x_{[-n,n]}). \end{aligned} \tag{5.6}$$

Cylinder sets are clopen sets, i.e. they are open and closed at the same time. As basis of the topology cylinder sets are open. To see that cylinder sets are closed, observe that the complement is a countable union of cylinder sets and hence open.

The metric from (5.5) captures the idea that points of a shift space are close when large central blocks of their coordinates agree. A sequence of points $(x^{(n)})_{n\in\mathbb{N}}$ in a shift space X converges exactly when, for each $i \geq 0$, the central $(2i+1)$-blocks stabilise, i.e. $x^{(n)} \longrightarrow x$ if and only if, for each $i \geq 0$, there is an n_i such that

$$x^{(n)}_{[i,-i]} = x_{[i,-i]}$$

for all $n \geq n_i$.

The shift map $\sigma_X : X \longrightarrow X$ is continuous, since the following holds: If two points in X are close, they agree on a large central block, hence their images agree on a large central block just shifted one place to the left.

By [**54**, Theorem 6.1.21] we know that a subset of $\mathbf{A}^{\mathbb{Z}}$ is a shift space if and only if it is shift-invariant and compact. With the metric d from (5.5) a shiftspace $X \subset \mathbf{A}^{\mathbb{Z}}$ is a compact metric space, and hence a compact

Hausdorff space. Therefore, the shift map σ_X together with X form an invertible dynamical system, referred to as *shift dynamical system.*

5.1.4. Shift Dynamical Systems Arising from Substitutions. Let $\mathbf{A}$ be a finite alphabet. A morphism $\varrho : \mathbf{A}^* \to \mathbf{A}^*$ induces a map from $\mathbf{A}^{\mathbb{Z}}$ to $\mathbf{A}^{\mathbb{Z}}$, which we also call ϱ, by

$$\varrho(x) := \ldots \varrho(x_{-1}).\varrho(x_0)\varrho(x_1)\ldots.$$

The map ϱ is referred to as *substitution* on $\mathbf{A}$. Note that $\varrho : \mathbf{A}^{\mathbb{Z}} \to \mathbf{A}^{\mathbb{Z}}$ is continuous.

A substitution ϱ on $\mathbf{A}$ is called *irreducible*, if for every pair of $a, b \in \mathbf{A}$ there is an $n = n(a, b) \in \mathbb{N}$ such that $\varrho^n(a)$ contains b. The substitution ϱ is said to be *primitive*, if there exits an $n \in \mathbb{N}$ such that $\varrho^n(a)$ contains b for every $a, b \in \mathbf{A}$. Note that, in this case, n can be chosen independently of a and b.

A finite word $w \in \mathbf{A}^*$ is called *legal* for the primitive substitution ϱ on $\mathbf{A}^{\mathbb{Z}}$, if it occurs as a factor of $\varrho^n(a)$ for some letter $a \in \mathbf{A}$ and $n \in \mathbb{N}$. Legal words have the property that they are mapped to legal words under the substitution.

A point $w \in \mathbf{A}^{\mathbb{Z}}$ is called a *fixed point* of a primitive substitution ϱ on $\mathbf{A}^{\mathbb{Z}}$ if $\varrho(w) = w$ and $w_{-1}.w_0$ is a legal two-letter word of ϱ.

Proposition 5.4 ([**7**]). *Let ϱ be a primitive substitution on the alphabet $\mathbf{A}$ with $|\mathbf{A}| \geq 2$. Then there exists a point $w \in \mathbf{A}^{\mathbb{Z}}$ and $n \geq 1$ such that $w = \varrho^n(w)$, i.e. w is a fixed point of ϱ^n.* □

Corollary 5.5. *If ϱ is a primitive substitution on the alphabet $\mathbf{A}$ with $|\mathbf{A}| \geq 2$, then the following holds:*

(i) $\lim_{n \longrightarrow \infty} |\varrho^n(a)| = \infty$ *for every letter $a \in \mathbf{A}$.*

(ii) *There exists a letter $a \in \mathbf{A}$ and $n \in \mathbb{N}$ so that $\varrho^n(a)$ begins with a.*

□

EXAMPLE 5.6. *As already mentioned in the introduction, the* Thue-Morse substitution *is a substitution on the alphabet* $\mathbf{A}_2 = \{0, 1\}$ *defined by*

$$t : \begin{array}{l} 0 \mapsto 01 \\ \\ 1 \mapsto 10. \end{array}$$

Two distinct bi-infinite fixed points w *and* $\widetilde{w}$ *are obtained as limits of the iteration of* t^2*,*

$$0.0 \overset{t^2}{\longmapsto} 0110.0110 \overset{t^2}{\longmapsto} 01101001100101 10.01101001100101 10 \overset{t^2}{\longmapsto} \ldots \overset{t^2}{\longmapsto} = w$$

$$1.0 \overset{t^2}{\longmapsto} 1001.0110 \overset{t^2}{\longmapsto} 10010110011010 01.01101001100101 10 \overset{t^2}{\longmapsto} \ldots \overset{t^2}{\longmapsto} = \widetilde{w}$$

EXAMPLE 5.7. *Another famous substitution on* $\mathbf{A}_2 = \{0, 1\}$ *is defined by*

$$f : \begin{array}{l} 0 \mapsto 01 \\ \\ 1 \mapsto 0 \end{array}$$

and is called the Fibonacci substitution. *Here, two distinct bi-infinite fixed points* w *and* $\widetilde{w}$ *are again obtained as limits of the iteration of* f^2*,*

$$\begin{array}{llllll} 0.0 & \overset{f^2}{\longmapsto} 010.010 & \overset{f^2}{\longmapsto} 01001010.0100101001 & \overset{f^2}{\longmapsto} \ldots & \overset{f^2}{\longmapsto} & = w \\ 1.0 & \overset{f^2}{\longmapsto} 01.010 & \overset{f^2}{\longmapsto} 01001.01001010 & \overset{f^2}{\longmapsto} \ldots & \overset{f^2}{\longmapsto} & = \widetilde{w} \end{array}$$

Let $w \in \mathbf{A}^{\mathbb{Z}}$ be a fixed point of the primitive substitution ϱ on $\mathbf{A}$. Define the *closed orbit* of w in $\mathbf{A}^{\mathbb{Z}}$ as

$$O(w) := \overline{\{\sigma^n(w) \mid n \in \mathbb{Z}\}}.$$

By definition $O(w)$ is a closed and shift-invariant subset of the compact metric space $\mathbf{A}^{\mathbb{Z}}$. Therefore $O(w)$ is a shift-space, see [**54**, Theorem 6.1.21], and $(O(w), \sigma_{O(w)})$ is a shift dynamical system; .

Now, we cite a result from [**72**, p. 105], which directly implies that the entropy of the shift space $O(w)$ vanishes.

PROPOSITION 5.8. *Let* w *be the fixed point of some primitive substitution* ϱ *on* $\mathbf{A}$ *and let* $c(n)$ *be the complexity function of* $O(w)$*. Then there exists*

a positive constant κ *such that*

$$c(n) \leq \kappa\, n \qquad \textit{for every } n \geq 1.$$

COROLLARY 5.9. *Let* w *be the fixed point of some primitive substitution* ϱ *on* $\mathbf{A}$ *and let* $c(n)$ *be the complexity function of* $O(w)$*. Then*

$$h(O(w)) = 0.$$

5.2. Topological Entropy

In this section we only state most of the results. For their proofs see for example [**86**, Chapter 7] and [**70**, Section 6.3].

5.2.1. Open Covers. Let X be a non-empty compact topological Hausdorff space which is not empty. We are interested in open covers of X which will be denoted by $\mathscr{U}$ or $\mathscr{V}$. If $\mathscr{U}$ and $\mathscr{V}$ are open covers of X their *join* $\mathscr{U} \vee \mathscr{V}$ is the open cover which consists of all sets of the form $U \cap V$ where $U \in \mathscr{U}$ and $V \in \mathscr{V}$. Similarly we define the join $\bigvee_{i=1}^{n} \mathscr{U}_i$ of any finite set of open covers of X.

An open cover $\mathscr{U}$ of X is a *refinement* of an open cover $\mathscr{V}$, written as $\mathscr{U} \geq \mathscr{V}$, if every member of $U \in \mathscr{U}$ is a subset of some $V \in \mathscr{V}$. Hence $\mathscr{U} \vee \mathscr{V} \geq \mathscr{U}$ for any open covers $\mathscr{U}, \mathscr{V}$. Note that if $\mathscr{V}$ is a subcover of $\mathscr{U}$ then $\mathscr{V} \geq \mathscr{U}$.

5.2.2. Entropy of an Open Cover. If $\mathscr{U}$ is an open cover of X let $N(\mathscr{U})$ be the number of sets in a finite subcover of X with smallest cardinality. The *entropy* of $\mathscr{U}$ is defined by

$$H(\mathscr{U}) := \log N(\mathscr{U}).$$

Note that

(i) $H(\mathscr{U}) \geq 0$

(ii) $H(\mathscr{U}) = 0$ if and only if $N(\mathscr{U}) = 1$ if and only if $X \in \mathscr{U}$.

(iii) If $\mathscr{V} \leq \mathscr{U}$ then $H(\mathscr{V}) \leq H(\mathscr{U})$.

(iv) $H(\mathscr{U} \vee \mathscr{V}) \leq H(\mathscr{U}) + H(\mathscr{V})$

5.2.3. Entropy of Continuous Maps. If $\mathscr{U}$ is an open cover of X and
$T : X \to X$ is continuous then $T^{-1}\mathscr{U}$ is the open cover of X which consists of all sets $T^{-1}U$ where $U \in \mathscr{U}$. Note that

(i) $T^{-1}(\mathscr{U} \vee \mathscr{V}) = T^{-1}(\mathscr{U}) \vee T^{-1}(\mathscr{V})$

(ii) $\mathscr{U} > \mathscr{V}$ implies $T^{-1}(\mathscr{U}) \geq T^{-1}(\mathscr{V})$.

(iii) $H(\mathscr{U}) \geq H(T^{-1}\mathscr{U})$. If T is also surjective then $H(\mathscr{U}) = H(T^{-1}\mathscr{U})$.

For $-\infty < m \leq n < \infty$, we define

$$\mathscr{U}_m^n := \bigvee_{i=m}^{n} T^{-i}\mathscr{U} = T^{-m}\mathscr{U} \vee T^{-(m+1)}\mathscr{U} \vee \ldots \vee T^{-n}\mathscr{U}.$$

THEOREM 5.10. *Let X be a non-empty compact topological Hausdorff space. If $\mathscr{U}$ is an open cover of X and $T : X \to X$ is continuous then*

$$h(\mathscr{U}, T) := \lim_{n \longrightarrow \infty} \frac{1}{n} H(\mathscr{U}_0^{n-1})$$

exists, and $h(\mathscr{U}, T) \leq H(\mathscr{U})$. □

DEFINITION 5.11. *Let X be a non-empty compact topological Hausdorff space. If $T : X \to X$ is continuous, the* topological entropy of *T is given by*

$$h_{\text{top}}(T) = \sup_{\mathscr{U}} h(\mathscr{U}, T)$$

where $\mathscr{U}$ ranges over all open covers of X.

Note that $0 \leq h_{\text{top}}(T) \leq \infty$ and that it is sufficient to take the supremum over finite open covers of X.

The following proposition provides a possibility to calculate topological entropy.

PROPOSITION 5.12. *Let X be a non-empty compact topological Hausdorff space and let $T : X \to X$ be continuous. If $\{\mathscr{U}_n\}_{n \in \mathbb{N}}$ is a refining sequence*

of open covers, i.e. $\mathscr{U}_1 \leq \mathscr{U}_2 \leq \dots$ *and for every finite open cover* $\mathscr{V}$ *of* X *there is an* $n \in \mathbb{N}$ *such that* $\mathscr{V} \leq \mathscr{U}_n$*, then*

$$h_{\text{top}}(T) = \lim_{n\to\infty} h(\mathscr{U}_n, T).$$

□

We conclude this chapter with the following theorem which tells us that for a shift space the topological entropy of its shift map and the combinatorial entropy of Definition 4.12 coincide.

THEOREM 5.13. *Let* $X \subset \mathbf{A}^{\mathbb{Z}}$ *be a shift space and let* $\sigma_X = \sigma$ *be its shift map. The topological entropy of* σ_X *equals the combinatorial entropy of* X*, i.e.*

$$h_{\text{top}}(\sigma_X) = h(X).$$

PROOF. For $i \in \mathbb{N}_0$ define

$$\mathscr{U}_i = \{C_{-i}(u) \mid u \in \mathcal{B}_{2i+1}(X)\},$$

where $C_{-i}(u) = C^X_{-i}(u)$ are cylinder sets. Obviously, $\{\mathscr{U}_i\}_{i\in\mathbb{N}}$ is a refining sequence of open covers of X.
For $-\infty < m \leq n < \infty$ the following holds:

$$\begin{aligned}(\mathscr{U}_i)^n_m = \bigvee_{j=m}^{n} \sigma^{-j}\mathscr{U}_i &= \{C_{-i+m}(u_m) \cap \dots \cap C_{-i+n}(u_n) \mid u_j \in \mathcal{B}_{2i+1}(X)\} \\ &= \{C_{-i+m}(u) \mid u \in \mathcal{B}_{2i+n-m+1}(X)\}\end{aligned}$$

as $|u| = i + n - (-i + m) + 1 = 2i + n - m + 1$. With $\mathscr{U} := \mathscr{U}_0$ this implies that $\mathscr{U}_i = (\mathscr{U}_0)^i_{-i} = \mathscr{U}^i_{-i}$ and

$$(\mathscr{U}_i)^{n-1}_0 = \{C_{-i}(u) \mid u \in \mathcal{B}_{2i+n}(X)\} = \mathscr{U}^{i+n-1}_{-i}. \tag{5.7}$$

A comparison of $\mathscr{U}^{n-1}_0$ and $(\mathscr{U}_i)^{n-1}_0$ leads to the relation $H(\mathscr{U}^{n-1}_0) \leq H((\mathscr{U}_i)^{n-1}_0)$ which implies that $h(\mathscr{U}, \sigma) \leq h(\mathscr{U}_i, \sigma)$. Hence for all $i \in \mathbb{N}$

we infer with (5.7) that

$$\begin{aligned} h(\mathscr{U},\sigma) \le h(\mathscr{U}_i,\sigma) &= \lim_{n\to\infty}\frac{1}{n}H((\mathscr{U}_i)_0^{n-1}) \\ &= \lim_{n\to\infty}\frac{1}{n}H(\mathscr{U}_{-i}^{i+n-1}) = \lim_{n\to\infty}\frac{1}{n}H(\mathscr{U}_i^{i+n-1}\vee\mathscr{U}_{-i}^{i-1}) \\ &\le \lim_{n\to\infty}\frac{1}{n}(H(\mathscr{U}_i^{i+n-1})+H(\mathscr{U}_{-i}^{i-1})) = \lim_{n\to\infty}\frac{1}{n}H(\mathscr{U}_0^{n-1}) \\ &= h(\mathscr{U},\sigma), \end{aligned}$$

which implies that $h(\mathscr{U},\sigma) = h(\mathscr{U}_i,\sigma)$ for all $i \in \mathbb{N}$. As at least $|\mathcal{B}_n(X)|$ elements of the open cover $\mathscr{U}_0^{n-1} = \{C_0(u) \mid u \in \mathcal{B}_n(X)\}$ are needed to cover X, we infer that

$$h_{\mathrm{top}}(\sigma) = \lim_{n\to\infty} h(\mathscr{U}_n,\sigma) = h(\mathscr{U},\sigma) = \lim_{n\to\infty}\frac{1}{n}H(\mathscr{U}_0^{n-1}) = h(X).$$

□

CHAPTER 6

The Lower Bound for the Entropy

In this chapter we prove two theorems which provide a procedure to calculate lower bounds for the entropy of integer and rational powerfree words. We generalise and explain in detail Kolpakov's ideas from [**49**], which start with a Perron-Frobenius argument and lead, via several inductive steps, to an estimate of the number of certain power containing words. This estimation results in a procedure to calculate lower bounds for the entropy of integer powerfree words as well as rational powerfree words.

We start by pointing out the basic idea of the proofs and proceed with a reformulation of what we have to show. For technical reasons, we treat integer and rational powerfree words separately.

Before we start off let us simplify some existing and set up some new notation. From now on, if not specified otherwise, let $\mathcal{F}$ stand for an infinite set of k-powerfree or α^+-powerfree words over an alphabet $\mathbf{A}_\ell$, where $k, \ell \in \mathbb{N}$ and $\alpha \in \mathbb{Q}$. A word $w \in \mathcal{F}$ is called *powerfree* in both cases and it will be clear from the context what is meant. We will repeatedly need the following sets:

$$\text{(6.1)}\quad \mathcal{L}_m := \mathcal{L}_m(\mathbf{A}_\ell) = \{v \in \mathbf{A}_\ell^* \mid \text{every factor } u \text{ of } v \text{ with } |u| = m \text{ is open or } u \notin \mathcal{F}(m)\}$$

$$\text{(6.2)}\quad \mathcal{F}_m := \mathcal{F}_m(\mathbf{A}_\ell) = \mathcal{L}_m \cap \mathcal{F}$$

$$\text{(6.3)}\quad \mathcal{F}_m^{(w)}(n) := \{v \in \mathcal{F}_m \mid |v| = n, v = xw\}.$$

Recall from (4.6) and (4.11) that $\mathcal{F}'(m)$ stands for the set of equivalence classes of powerfree words of length m and that $\mathcal{F}''(m) = \{w_1, \ldots, w_s\}$ denotes the set of open words among them.

For every real number β we define

$$\lfloor \beta \rfloor := \max \{n \in \mathbb{Z} \mid n \leq \beta\} \quad \text{and} \quad \lceil \beta \rceil := \min \{n \in \mathbb{Z} \mid n \geq \beta\}. \tag{6.4}$$

Repeatedly we will need the concept of multisets. A *multiset* is a set with repeated elements. For example $[0, 0, 1, 2, 2, 2]$ is a multiset. More precisely, a *finite multiset* M on a set S with $|S| = n$ elements is a function $\nu : S \to \mathbb{N}$ such that $\sum_{x \in S} \nu(x) < \infty$. One regards $\nu(x)$ as the number of repetitions of x and we write

$$M = [\underbrace{x_1, \ldots, x_1}_{\nu(x_1) \text{ times}}, \ldots, \underbrace{x_n, \ldots, x_n}_{\nu(x_n) \text{ times}}].$$

Of course the union as well as the intersection of multisets is again a multiset; see [**80**] for details. Note that $[0, 0, 1, 2, 2, 2]$ stands for a multiset whereas $\{0, 0, 1, 2, 2, 2\} = \{0, 1, 2\}$ continues to denote a 'normal' set.

6.1. Strategy

6.1.1. Idea. The derivation of the lower bound for the entropy $h(\mathcal{F})$ is based on the following idea. Assume that Δ_m from Definition 4.18 is irreducible. The Perron-Frobenius Theorem for irreducible matrices, see for example [**54**, Theorem 4.2.3], tells us that the largest eigenvalue λ_m, called the Perron eigenvalue of Δ_m, possess a strictly positive right eigenvector $(x_1, \ldots, x_s)$, which is unique up to a positive multiple. Moreover, we assume that $\lambda_m > 1$ and normalise the eigenvector according to $\sum_{i=1}^{s} x_i = 1$.

For $w_i \in \mathcal{F}''(m)$ we define

$$d_m(n) := \sum_{i=1}^{s} x_i |\mathcal{F}_m^{(w_i)}(n)| \tag{6.5}$$

and note that

$$c_{\mathcal{F}}(n) \geq \sum_{i=1}^{s} |\mathcal{F}^{(w_i)}(n)| \geq \sum_{i=1}^{s} x_i |\mathcal{F}_m^{(w_i)}(n)| = d_m(n).$$

Moreover, $d_m(m) = 1$ since by definition $|\mathcal{F}_m^{(w_i)}(m)| = 1$.

The idea of this chapter is to show that for some $\gamma \in (1, \lambda_m)$ and every $n \geq m$ the inequality

$$d_m(n+1) \geq \gamma d_m(n) \tag{6.6}$$

holds. We conclude that

$$d_m(n) \geq \gamma^{n-m}$$

and hence

$$h(\mathcal{F}) = \lim_{n\to\infty} \tfrac{1}{n} \log(c_{\mathcal{F}}(n)) \geq \lim_{n\to\infty} \tfrac{1}{n} \log(d_m(n)) \geq \log(\gamma) - \lim_{n\to\infty} \tfrac{m}{n} \log(\gamma) = \log(\gamma),$$

thus providing the lower bound $\log(\gamma)$ for the entropy $h(\mathcal{F})$.

6.1.2. Reformulation. The main task is to prove inequality (6.6). The reformulation of this task requires the definition of the following sets.

For any $w_i \in \mathcal{F}''(m)$ and $n \geq m$ we set

(6.7)
$$\mathcal{G}^{(w_i)}(n+1) := \{w \in \mathcal{L}_m^{(w_i)}(n+1) \mid w[1:n], w[n-m+1, n+1] \in \mathcal{F}\},$$

(6.8)
$$\mathcal{H}^{(w_i)}(n+1) := \{w \in \mathcal{G}^{(w_i)}(n+1) \mid w = uv \text{ where } v \notin \mathcal{F}\}.$$

Note that for $w \in \mathcal{G}^{(w_i)}(n+1)$, the powerfree prefix $w[1:n]$ and the powerfree suffix $w[n-m+1, n+1]$ overlap on $w[n-m+1:n]$, which is an ancestor of w_i. Considering (6.3) makes it obvious that

$$|\mathcal{F}_m^{(w_i)}(n+1)| = |\mathcal{G}^{(w_i)}(n+1)| - |\mathcal{H}^{(w_i)}(n+1)|. \tag{6.9}$$

and hence we have

$$d_m(n+1) = \sum_{i=1}^{s} x_i |\mathcal{G}_m^{(w_i)}(n+1)| - \sum_{i=1}^{s} x_i |\mathcal{H}_m^{(w_i)}(n+1)|. \tag{6.10}$$

For $1 \leq i \leq s$ we define

$$\mathcal{Q}(i) := \{w \in \mathcal{F}'(m) \mid w \text{ is a quasi-ancestor of } w_i\}. \tag{6.11}$$

If $u, v \in \mathbf{A}_\ell^*$ are isomorphic we have $|\mathcal{F}_m^{(u)}(n)| = |\mathcal{F}_m^{(v)}(n)|$ which implies that

$$|\mathcal{G}^{(w_i)}(n+1)| = \sum_{w \in \mathcal{Q}(i)} |\mathcal{F}_m^{(w)}(n)|. \tag{6.12}$$

With Definition 4.18 we infer that

$$\begin{aligned}
\sum_{i=1}^{s} x_i |\mathcal{G}^{(w_i)}(n+1)| &= \sum_{i=1}^{s} (x_i \sum_{w \in \mathcal{Q}(i)} |\mathcal{F}_m^{(w)}(n)|) \\
&= (x_1, \ldots, x_s) \begin{pmatrix} \delta_{11} & \delta_{21} & \ldots & \delta_{s1} \\ \delta_{12} & \delta_{22} & \ldots & \delta_{s2} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{1s} & \delta_{2s} & \ldots & \delta_{ss} \end{pmatrix} \begin{pmatrix} |\mathcal{F}_m^{(w_1)}(n)| \\ |\mathcal{F}_m^{(w_2)}(n)| \\ \vdots \\ |\mathcal{F}_m^{(w_s)}(n)| \end{pmatrix} \\
&= (x_1, \ldots, x_s)\ \Delta_m^T\ (|\mathcal{F}_m^{(w_1)}(n)|, |\mathcal{F}_m^{(w_2)}(n)|, \ldots, \mathcal{F}_m^{(w_s)}(n)|)^T \\
&= \lambda_m(x_1, \ldots, x_s) \quad (|\mathcal{F}_m^{(w_1)}(n)|, |\mathcal{F}_m^{(w_2)}(n)|, \ldots, \mathcal{F}_m^{(w_s)}(n)|)^T \\
&= \lambda_m\ d_m(n).
\end{aligned} \tag{6.13}$$

Since $\mathcal{H}^{(w_i)}(m+1) = \varnothing$ and $d_m(m) = 1$ due to their definitions, we infer that

$$d_m(m+1) = \lambda_m. \tag{6.14}$$

Considering (6.10) and (6.13), we see that

$$d_m(n+1) = \lambda_m d_m(n) - \sum_{i=1}^{s} x_i |\mathcal{H}_m^{(w_i)}(n+1)|. \tag{6.15}$$

Hence, in order to show (6.6), we have to estimate $\sum_{i=1}^{s} x_i |\mathcal{H}^{(w_i)}(n+1)|$ by $d_m(n)$.

We now concentrate on the case of powerfree words for integer powers, and consider the case of rational powers in Section 6.3.

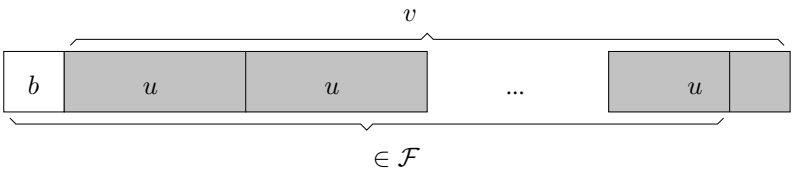


FIGURE 6.1. $v_t \in V_j^{(w_i)}$and $b \in B_t$ with $bv_t = bu^k$.

6.2. Words Avoiding Integer Powers

Let in this section $\mathcal{F} = \mathcal{F}^{(k)}(\mathbf{A}_\ell)$, where $k \in \mathbb{N}$ and $\mathbf{A}_\ell$ is an alphabet with $\ell \geq 2$ letters such that $\mathcal{F}$ is an infinite set. Moreover, let $m, p, q \in \mathbb{N}$ such that $2 < m \leq p$ and $kp - m \leq q$.

6.2.1. Polynomial. For the derivation of the lower bound we need to define a particular polynomial. The proof of Theorem 6.1 below will provide the motivation for its definition. We set

$$J := \left\{ \lfloor \tfrac{(m+1)}{k} \rfloor + 1, \ldots, \lfloor \tfrac{q+m}{k} \rfloor \right\}. \tag{6.16}$$

For every $j \in J$ and $w_i \in \mathcal{F}''(m)$ we define the set of all k-powers of minimal period j, which do not have closed words of length m as factors:

$$\begin{aligned} V_j^{(w_i)} :&= \left\{ v \in \mathcal{L}_m^{(w_i)} \mid v = u^k \text{ with } u \in \mathcal{F}(j), \operatorname{per}(v) = j \right\} \\ &= \{v_1, \ldots, v_o\}. \end{aligned} \tag{6.17}$$

Now, we define for $1 \leq t \leq o$

$$B_t := \{ b \in \mathbf{A}_\ell \mid bv_t[1 : kj-1] \in \mathcal{F}_m \}. \tag{6.18}$$

Note that since $bv[1 : kj-1]$ is k-powerfree we know that $B_t \subset \mathbf{A}_\ell \setminus \{v_t[j]\}$, see Figure 6.1 for an illustration. For $j \in J, w_i \in \mathcal{F}''(m)$ and $1 \leq t \leq o$ we define the following multiset

$$\begin{aligned} U_j(w_i) &:= \bigcup_{1 \leq t \leq o} U_{j,t}(w_i), \text{ where} \\ U_{j,t}(w_i) &:= \left[(bv_t[1 : m-1])' \mid b \in B_t \right] \subset \mathcal{F}''(m). \end{aligned} \tag{6.19}$$

Let $\eta_r(j) := \sum_{i=1}^{s} x_i e_j^{(r)}(w_i)$, where $e_j^{(r)}(w_i) := |\{w_r \in U_j(w_i)\}|$. We substitute $g(j) := kj - m$ and define

$$g(\lfloor \tfrac{(m+1)}{k} \rfloor + 1) = k\lfloor \tfrac{(m+1)}{k} \rfloor + k - m =: g_0$$

Furthermore, we set for $g_0 \leq g \leq q$

$$\eta_r'(g) := \begin{cases} \eta_r(\frac{g+m}{k}), & \text{if } g+m \text{ is divisible by } k, \\ 0, & \text{otherwise} \end{cases}.$$

Now, we are finally in a position to define inductively our polynomial

$$\mathcal{P}_m^{(p,q)}(x) := \sum_{g=g_0}^{q} \varrho_g \, x^g. \tag{6.20}$$

Recall that we denote by $(x_1, \ldots, x_s)$ the right strictly positive eigenvector corresponding to the Perron eigenvalue λ_m of Δ_m. We set

$$\varrho_{g_0} := \min_{1 \leq r \leq s} (\frac{\eta_r'(g_0)}{x_r}) \tag{6.21}$$

and $\nu_r := \eta_r'(g_0) - \varrho_{g_0} x_r$ for $r \in \{1, \ldots, s\}$. Let $\nu := (\nu_1, \ldots, \nu_s), \nu' = \Delta_m \nu = (\nu_1', \ldots, \nu_s')$ and finally $\eta_r''(g_0+1) := \eta_h'(g_0+1) + \nu_k'$.

Assume now that for some $g-1$ with $g_0 \leq g-1 < q-1$ we have already computed the numbers $\varrho_{g_0}, \ldots, \varrho_{g-1}$ and $\eta_1''(g), \ldots, \eta_s''(g)$. Then we define

$$\varrho_g := \min_{1 \leq r \leq s} (\tfrac{\eta_r''(g)}{x_r}) \tag{6.22}$$

so that all entries in $\widetilde{\nu} := (\eta_1''(g) - \varrho_g x_1, \ldots, \eta_s''(g) - \varrho_g x_s)$ are non-negative. Let $\widetilde{\nu}' := \Delta_m \widetilde{\nu} = (\widetilde{\nu_1}', \ldots, \widetilde{\nu_s}')$ and set for $r \in \{1, \ldots, s\}$

$$\eta_r''(g+1) := \eta_r'(g+1) + \widetilde{\nu_r}'.$$

For $g = q$, we define

$$\varrho_q := \max_{1 \leq r \leq s} (\tfrac{\eta_r''(q)}{x_r}). \tag{6.23}$$

Note that these definitions of the coefficients of the polynomial $\mathcal{P}_m^{(p,q)}(x)$ ensure that it does not matter how the eigenvector $(x_1, \ldots, x_r)$ is normalised.

6.2.2. The Lower Bound. The main result of this chapter is the following theorem. It is a generalisation of the method introduced in [**49**] and basically gives a procedure to calculate lower bounds for the entropy of integer powerfree words.

THEOREM 6.1. *Let $\mathbf{A}_\ell$ be a finite alphabet with $\ell \geq 2$ letters and let $k \in \mathbb{N}$ such that $\mathcal{F} = \mathcal{F}^{(k)}(\mathbf{A}_\ell)$ is an infinite set. For $m > 2$ let Δ_m be the corresponding matrix from Definition 4.18. Let Δ_m be irreducible and denote its Perron eigenvalue by λ_m. Moreover, for $p, q \in \mathbb{N}$ with $p \geq m$ and $q \geq kp - m$ construct the polynomial $\mathcal{P}_m^{(p,q)}$ of (6.20). If for some $\gamma \in (1, \lambda_m)$ and all $n \in \{m, \ldots, q + m - 1\}$.*

$$d_m(n+1) \geq \gamma d_m(n) \tag{6.24}$$

as well as

$$\lambda_m - \mathcal{P}_m^{(p,q)}(\tfrac{1}{\gamma}) - \frac{1}{\gamma^{(k-1)p-1}(\gamma^{(k-1)} - 1)} \geq \gamma \tag{6.25}$$

then

$$h(\mathcal{F}) \geq \log(\gamma).$$

PROOF. Recall that we intend to show that the inequality (6.24) holds for every $n \in \mathbb{N}$ since this gives us the lower bound $\log(\gamma)$ for the entropy $h(\mathcal{F})$. We argue by induction on $n \geq q + m \geq kp$. The basis is given by assumption (6.24). So assume that for $m \leq i \leq n - 1$ the inequality

$$d_m(i+1) \geq \gamma d_m(i)$$

holds. This implies directly that for $m \leq i \leq n - 1$

$$\frac{d_m(n)}{\gamma^{n-i}} \geq d_m(i). \tag{6.26}$$

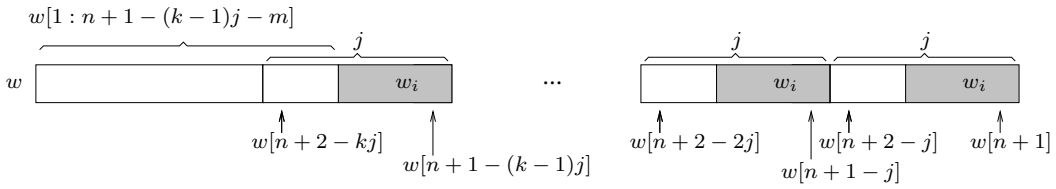


FIGURE 6.2. $w \in \mathcal{H}_j(w_i)$ where $j \in J_1$.

Recall from Section 6.1.2 that we have to estimate $\sum_{i=1}^{s} x_i |\mathcal{H}^{(w_i)}(n+1)|$ by $d_m(n)$. For any word $w \in \mathcal{H}^{(w_i)}(n+1)$ we can find the shortest suffix of w which is a k-power. Let v_w be this suffix. Since $|w| = n+1$, $\mathrm{per}(v_w)$ can at most be $\lfloor \frac{(n+1)}{k} \rfloor$ and because $|w[n-m+1 : n+1]| = m+1$, we know that $\mathrm{per}(v_w) > \lfloor \frac{(m+1)}{k} \rfloor$. Hence we have $\lfloor \frac{(m+1)}{k} \rfloor < \mathrm{per}(v_w) \leq \lfloor \frac{(n+1)}{k} \rfloor$. We define

$$\mathcal{H}_j^{(w_i)}(n+1) := \left\{ w \in \mathcal{H}^{(w_i)}(n+1) \mid \mathrm{per}(v_w) = j \right\} \tag{6.27}$$

for $\lfloor \frac{(m+1)}{k} \rfloor < j \leq \lfloor \frac{(n+1)}{k} \rfloor$, so that we can write

$$\mathcal{H}^{(w_i)}(n+1) = \sum_{\lfloor \frac{(m+1)}{k} \rfloor < j \leq \lfloor \frac{(n+1)}{k} \rfloor} \mathcal{H}_j^{(w_i)}(n+1). \tag{6.28}$$

Now, we distinguish

(1) $J_1 := \left\{ j \in \mathbb{N} \mid p < j \leq \left\lfloor \frac{(n+1)}{k} \right\rfloor \right\}$ and

(2) $J_2 := \left\{ j \in \mathbb{N} \mid \left\lfloor \frac{(m+1)}{k} \right\rfloor < j \leq p \right\}$.

It would be possible to take $p = m$, but our estimation in case (2) is often better and so it makes sense to exploit it as far as computationally possible. Note that $J_2 \subset J$ from (6.16).

(1) We start by considering J_1. Let $w \in \mathcal{H}_j^{(w_i)}(n+1)$ with $j \in J_1$. We know that $j > m$ and hence

$$\begin{aligned} w[n+1-kj+1 : n+1-(k-1)j] \\ &= w[n+1-(k-1)j+1 : n+1-(k-2)j] \\ &= \dots \\ &= w[n+1-j+1 : n+1], \end{aligned}$$

compare Figure 6.2.

This means that w_i is not only a suffix of w, but also occurs in particular at

$$w[n+1-(k-1)j-m+1 : n+1-(k-1)j] = w_i.$$

So w is determined uniquely by the prefix

$$w[1 : n+1-(k-1)j-m]$$

which implies that

$$|\mathcal{H}_j^{(w_i)}(n+1)| \leq |\mathcal{F}_m^{(w_i)}(n+1-(k-1)j)|. \tag{6.29}$$

Considering (6.26) this means

$$
\begin{aligned}
&\sum_{i=1}^{s} x_i \Big(\sum_{j \in J_1} |\mathcal{H}_j^{(w_i)}(n+1)| \Big) \\
&\qquad \leq \sum_{i=1}^{s} x_i \Big(\sum_{j \in J_1} |\mathcal{F}_m^{(w_i)}(n+1-(k-1)j)| \Big) \\
&\qquad = \sum_{j \in J_1} \Big(\sum_{i=1}^{s} x_i |\mathcal{F}_m^{(w_i)}(n+1-(k-1)j)| \Big) \\
&\qquad = \sum_{j \in J_1} d_m(n+1-(k-1)j) \\
&\qquad = \sum_{p < j \leq \lfloor (n+1)/k \rfloor} \frac{d_m(n)}{\gamma^{(k-1)j-1}} \\
&\qquad < \frac{d_m(n)}{\gamma^{(k-1)p-1}(\gamma^{(k-1)}-1)}.
\end{aligned}
\tag{6.30}
$$

(2) Now, we consider J_2, which is the more complicated case.

Let $w \in \mathcal{H}_j^{(w_i)}(n+1)$ with $j \in J_2$. Note that

$$w[n+1-kj+1 : n+1] = w[n-kj+2 : n+1]$$

is a k-power which has w_i as a suffix. It contains, by definition of $\mathcal{H}_j^{(w_i)}(n+1)$, no other k-powers as factors and we know, as $w \in \mathcal{L}_m^{(w_i)}(n+1)$, that every factor of length m is open. This means that $w[n-kj+2 : n+1] \in V_j^{(w_i)}$ and all possible k-powers of minimal period j are contained in $V_j^{(w_i)} = \{v_1, \ldots, v_o\}$, see (6.17). For $1 \leq t \leq o$ define

$$\mathcal{H}_{j,t}^{(w_i)}(n+1) := \Big\{ w \in \mathcal{H}_j^{(w_i)}(n+1) \mid w = y v_t \Big\}$$

and let $w \in \mathcal{H}_{j,t}^{(w_i)}(n+1)$. Note that $y \neq \varepsilon$, since we assumed that $j \leq p < \frac{(n+1)}{k}$.

Now, we follow the construction of the polynomial in Section 6.2.1. Let $u \in U_j(w_i)$, see (6.19), and note that w is determined uniquely by the prefix

$$w[1 : n+1-kj],$$

which is always a proper prefix since $j < \frac{(n+1)}{k}$ and hence $kj \leq n$, compare Figure 6.3 for an illustration. As $w[n+1-kj] \in B_t$, see (6.18), and $|u| = m$,

we have

$$|\mathcal{H}_{j,t}^{(w_i)}(n+1)| \leq \sum_{u \in U_{j,t}(w_i)} |\mathcal{F}_m^{(u)}(n-kj+m)|.$$

Thus

$$|\mathcal{H}_{j}^{(w_i)}(n+1)| = \sum_{k=1}^{t} |\mathcal{H}_{j,t}^{(w_i)}(n+1)| \leq \sum_{u \in U_{j}(w_i)} |\mathcal{F}_m^{(u)}(n-kj+m)|.$$

Moreover, we see that

(6.31)

$$\begin{aligned}
\sum_{i=1}^{s} x_i \sum_{j \in J_2} |\mathcal{H}_{j}^{(w_i)}(n+1)| &\leq \sum_{i=1}^{s} x_i \sum_{j \in J_2} \sum_{u \in U_j(w_i)} |\mathcal{F}_m^{(u)}(n-kj+m)| \\
&= \sum_{j \in J_2} \sum_{i=1}^{s} \sum_{u \in U_j(w_i)} x_i |\mathcal{F}_m^{(u)}(n-kj+m)| \\
&= \sum_{j \in J_2} \sum_{i=1}^{s} \sum_{r=1}^{s} x_i e_j^{(r)}(w_i)) |\mathcal{F}_m^{(w_r)}(n-kj+m)| \\
&= \sum_{j \in J_2} \sum_{r=1}^{s} (\sum_{i=1}^{s} x_i e_j^{(r)}(w_i)) |\mathcal{F}_m^{(w_r)}(n-kj+m)|) \\
&= \sum_{j \in J_2} \sum_{r=1}^{s} \eta_r(j) |\mathcal{F}_m^{(w_r)}(n-kj+m)|) \\
&\leq \sum_{g=g_0}^{q} \sum_{r=1}^{s} \eta'_h(g) |\mathcal{F}_m^{(w_r)}(n-g)|
\end{aligned}$$

We will show now that this sum is majorized by

$$\sum_{g=g_0}^{q} \varrho_g d_m(n-g).$$

Therefore, we first show by induction that for $\ell \in \{g_0, \ldots, q-1\}$.

(6.32)

$$\begin{aligned}
&\sum_{g=g_0}^{\ell+1} \sum_{r=1}^{s} \eta'_h(g) |\mathcal{F}_m^{(w_r)}(n-g)| \\
&\qquad \leq \sum_{r=1}^{s} \eta''_r(\ell+1) |\mathcal{F}_m^{(w_r)}(n-\ell-1)| + \sum_{g=g_0}^{\ell} \varrho_g d_m(n-g).
\end{aligned}$$

Note that in the proof of the inequality we only need that $n \geq \ell + 1$.

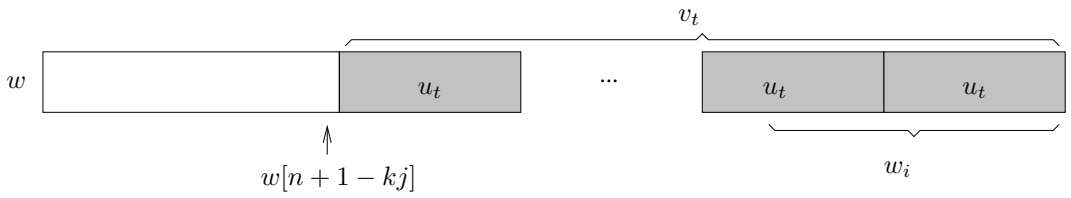


FIGURE 6.3. $w \in \mathcal{H}_{j,k}(w_i)$ where $j \in J_2$ and $v_t = u_t^k$.

Recall that we set $\varrho_{g_0} = \min_{1\le r\le s}(\frac{\eta'_r(g_0)}{x_r})$ and that the vectors

$$\nu = (\eta'_1(g_0) - \varrho_{g_0}x_1, \dots, \eta'_1(g_0) - \varrho_{g_0}x_1) = (\nu_1, \dots, \nu_s) \text{ and}$$
$$\nu' = \Delta_m\nu = (\nu'_1, \dots, \nu'_s)$$

are non-negative. Inserting these definitions gives

$$\begin{aligned}
&\sum_{r=1}^{s} \eta'_r(g_0)|\mathcal{F}_m^{(w_r)}(n-g_0)| \\
&= \varrho_{g_0} d_m(n-g_0) - \varrho_{g_0}\sum_{r=1}^{s} x_r|\mathcal{F}_m^{(w_r)}(n-g_0)| + \sum_{r=1}^{s}\eta'_r(g_0)|\mathcal{F}_m^{(w_r)}(n-g_0)| \\
&= \varrho_{g_0} d_m(n-g_0) + \sum_{r=1}^{s}(\eta'_r(g_0) - \varrho_{g_0}x_r)|\mathcal{F}_m^{(w_r)}(n-g_0)| \\
&= \varrho_{g_0} d_m(n-g_0) + \sum_{r=1}^{s}\nu_r|\mathcal{F}_m^{(w_r)}(n-g_0)|.
\end{aligned}$$

From (6.9) and (6.12) we know that for $r = 1, \dots, s$

$$|\mathcal{F}_m^{(w_r)}(n-g_0)| \le |\mathcal{G}^{(w_r)}(n-g_0)| = \sum_{w\in\mathcal{Q}(r)} |\mathcal{F}_m^{(w)}(n-g_0-1)|.$$

Thus

$$\begin{aligned}
\sum_{r=1}^{s}\nu_r|\mathcal{F}_m^{(w_r)}(n-g_0)| &\le \sum_{r=1}^{s}\nu_r|\mathcal{G}^{(w_r)}(n-g_0)| = \sum_{r=1}^{s}\nu_r \sum_{w\in\mathcal{Q}(r)} |\mathcal{F}_m^{(w)}(n-g_0-1)| \\
&= (\nu_1, \dots, \nu_s)\ \Delta_m^T\ (|\mathcal{F}_m^{(w_1)}(n-g_0-1)|, \dots, \mathcal{F}_m^{(w_s)}(n-g_0-1)|)^T \\
&= \sum_{r=1}^{s}\nu'_r|\mathcal{F}_m^{(w_r)}(n-g_0-1)|.
\end{aligned}$$

If we recall that $\eta''_r(g_0+1) = \eta'_r(g_0+1) + \nu'_r$ we can infer (6.32) for $\ell = g_0$:

$$\begin{aligned}
&\sum_{g=g_0}^{g_0+1}\sum_{r=1}^{s}\eta'_r(g)|\mathcal{F}_m^{(w_r)}(n-g)| \\
&= \sum_{r=1}^{s}\eta'_h(g_0)|\mathcal{F}_m^{(w_r)}(n-g_0)| + \sum_{r=1}^{s}\eta'_h(g_0+1)|\mathcal{F}_m^{(w_r)}(n-g_0-1)| \\
&= \varrho_{g_0}d_m(n-g_0) + \sum_{r=1}^{s}\nu_r|\mathcal{F}_m^{(w_r)}(n-g_0)| + \sum_{r=1}^{s}\eta'_h(g_0+1)|\mathcal{F}_m^{(w_r)}(n-g_0-1)| \\
&\le \varrho_{g_0}d_m(n-g_0) + \sum_{r=1}^{s}\nu'_r|\mathcal{F}_m^{(w_r)}(n-g_0-1)| + \sum_{r=1}^{s}\eta'_r(g_0+1)|\mathcal{F}_m^{(w_r)}(n-g_0-1)| \\
&\le \sum_{r=1}^{s}\eta''_r(g_0+1)|\mathcal{F}_m^{(w_r)}(n-g_0-1)| + \varrho_{g_0}d_m(n-g_0)
\end{aligned}$$

Assume now that for some $\ell-1$ with $g_0 \le \ell-1 < q-1$ we have already shown that (6.32) holds. Inserting the definitions gives

$$\begin{aligned}
&\sum_{r=1}^{s}\eta''_h(\ell)|\mathcal{F}_m^{(w_r)}(n-\ell)| \\
&= \varrho_\ell\, d_m(n-\ell) - \varrho_\ell\sum_{r=1}^{s}x_r|\mathcal{F}_m^{(w_r)}(n-\ell)| + \sum_{r=1}^{s}\eta''_h(\ell)|\mathcal{F}_m^{(w_r)}(n-\ell)| \\
&= \varrho_\ell\, d_m(n-\ell) + \sum_{r=1}^{s}(\eta''_r(\ell) - \varrho_\ell x_r)|\mathcal{F}_m^{(w_r)}(n-\ell)| \\
&= \varrho_\ell\, d_m(n-\ell) + \sum_{r=1}^{s}\widetilde{\nu}_r|\mathcal{F}_m^{(w_r)}(n-\ell)|
\end{aligned}$$

From (6.9) and (6.12) we know that for $h = 1, \ldots, s$

$$|\mathcal{F}_m^{(w_r)}(n-\ell) \le |\mathcal{G}^{(w_r)}(n-\ell)| = \sum_{w\in\mathcal{Q}(r)}|\mathcal{F}_m^{(w)}(n-\ell-1)|$$

and hence

$$\begin{aligned}
\sum_{r=1}^{s}\widetilde{\nu}_r|\mathcal{F}_m^{(w_r)}(n-\ell)| &\le \sum_{r=1}^{s}\widetilde{\nu}_r\sum_{w\in\mathcal{Q}(r)}|\mathcal{F}_m^{(w)}(n-\ell-1)| \\
&= (\widetilde{\nu}_1,\ldots,\widetilde{\nu}_s)\;\Delta_m^T\;(|\mathcal{F}_m^{(w_1)}(n-\ell-1)|,\ldots,\mathcal{F}_m^{(w_s)}(n-\ell-1)|)^T \\
&= \sum_{r=1}^{s}\widetilde{\nu}'_r|\mathcal{F}_m^{(w_r)}(n-\ell-1)|.
\end{aligned}$$

Thus we get

$$\begin{aligned}
&\sum_{r=1}^{s} \eta_h''(\ell)|\mathcal{F}_m^{(w_r)}(n-\ell)| + \sum_{r=1}^{s} \eta_h'(\ell+1)|\mathcal{F}_m^{(w_r)}(n-\ell-1)| \\
&= \varrho_\ell d_m(n-\ell) + \sum_{r=1}^{s} \widetilde{\nu}_h|\mathcal{F}_m^{(w_r)}(n-\ell)| + \sum_{r=1}^{s} \eta_r'(\ell+1)|\mathcal{F}_m^{(w_r)}(n-\ell-1)| \\
&\le \varrho_\ell d_m(n-\ell) + \sum_{r=1}^{s} \widetilde{\nu}_r'|\mathcal{F}_m^{(w_r)}(n-\ell-1)| + \sum_{r=1}^{s} \eta_r'(\ell+1)|\mathcal{F}_m^{(w_r)}(n-\ell-1)| \\
&= \sum_{r=1}^{s} \eta_h''(\ell+1)|\mathcal{F}_m^{(w_r)}(n-\ell-1)| + \varrho_\ell\, d_m(n-\ell),
\end{aligned}$$

which provides the induction step for (6.32):

$$\begin{aligned}
&\sum_{g=g_0}^{\ell+1} \sum_{r=1}^{s} \eta_r'(g)|\mathcal{F}_m^{(w_r)}(n-g)| \\
&= \sum_{g=g_0}^{\ell} \sum_{r=1}^{s} \eta_r'(g)|\mathcal{F}_m^{(w_r)}(n-g)| + \sum_{r=1}^{s} \eta_r'(\ell+1)|\mathcal{F}_m^{(w_r)}(n-\ell-1)| \\
&\le \sum_{r=1}^{s} \eta_r''(\ell)|\mathcal{F}_m^{(w_r)}(n-\ell)| + \sum_{g=g_0}^{\ell-1} \varrho_\ell d_m(n-g) + \sum_{r=1}^{s} \eta_r'(\ell+1)|\mathcal{F}_m^{(w_r)}(n-\ell-1)| \\
&= \sum_{r=1}^{s} (\eta_r''(\ell)|\mathcal{F}_m^{(w_r)}(n-\ell)| + \eta_r'(\ell+1)|\mathcal{F}_m^{(w_r)}(n-\ell-1)|) + \sum_{g=g_0}^{\ell-1} \varrho_\ell d_m(n-g) \\
&\le \sum_{r=1}^{s} \eta_r''(\ell+1)|\mathcal{F}_m^{(w_r)}(n-\ell-1)| + \varrho_g d_m(n-\ell) + \sum_{g=g_0}^{\ell-1} \varrho_\ell d_m(n-g)
\end{aligned}$$

Now we resume the estimation from (6.31) by applying (6.32) for $\ell = q-1$:

$$\sum_{g=g_0}^{q} \sum_{r=1}^{s} \eta_r'(g)|\mathcal{F}_m^{(w_r)}(n-g)| \le \sum_{h=1}^{s} \eta_h''(q)|\mathcal{F}_m^{(w_h)}(n-q)| + \sum_{g=d_0}^{q-1} \varrho_g d_m(n-g)$$

and

$$\begin{aligned}
&\sum_{r=1}^{s} \eta_h''(q)|\mathcal{F}_m^{(w_r)}(n-q)| \\
&= \varrho_q d_m(n-q) - \varrho_q \sum_{r=1}^{s} x_r|\mathcal{F}_m^{(w_r)}(n-q)| + \sum_{r=1}^{s} \eta_h''(q)|\mathcal{F}_m^{(w_r)}(n-q)| \\
&= \varrho_q d_m(n-q) + \sum_{r=1}^{s} (\eta_r''(q) - \varrho_q x_r)|\mathcal{F}_m^{(w_r)}(n-q)| \\
&\leq \varrho_q d_m(n-q).
\end{aligned}$$

So we have finally shown that

$$\begin{aligned}
\sum_{i=1}^{s} x_i \sum_{j\in J} |\mathcal{H}_j^{(w_i)}(n+1)| &\leq \sum_{g=g_0}^{q} \varrho_g d_m(n-g) \leq \sum_{g=g_0}^{q} \varrho_g \frac{d_m(n)}{\gamma^{n-(n-g)}} \\
&= d_m(n) \sum_{g=g_0}^{q} \frac{\varrho_g}{\gamma^g} = d_m(n)\mathcal{P}_m^{(p,q)}(\tfrac{1}{\gamma}).
\end{aligned} \tag{6.33}$$

Note that, since $g_0 = k\lfloor\frac{(m+1)}{k}\rfloor + k - m > k(\frac{m+1}{k} - 1) + k - m = m + 1 - k + k - m = 1$, we know that $n - g_0 < n - 1$.

The combination of the result (6.30) for J_1 and the result (6.33) for J_2 now implies that

$$\sum_{i=1}^{s} x_i|\mathcal{H}^{(w_i)}(n+1)| < d_m(n)\ (\mathcal{P}_m^{(p,q)}(\tfrac{1}{\gamma}) + \tfrac{1}{\gamma^{(k-1)p-1}(\gamma^{(k-1)}-1)}).$$

and hence

$$\begin{aligned}
d_m(n+1) &= \sum_{i=1}^{s} x_i|\mathcal{G}^{(w_i)}(n+1)| - \sum_{i=1}^{s} x_i|\mathcal{H}^{(w_i)}(n+1)| \\
&= \lambda_m d_m(n) - \sum_{i=1}^{s} x_i|\mathcal{H}^{(w_i)}(n+1)| \\
&> d_m(n)(\lambda_m - (\mathcal{P}_m^{(p,q)}(\tfrac{1}{\gamma}) + \tfrac{1}{\gamma^{(k-1)p-1}(\gamma^{(k-1)}-1)}) \\
&\geq \gamma d_m(n),
\end{aligned}$$

which completes the proof. □

It can take very long to check computationally that the assumption (6.24) holds. Alternatively, we can compute inductively $\tau_n \in \mathbb{R}$ for

$n \in \{m, \dots, q+m-1\}$, such that

$$d_m(n+1) \geq \tau_n d_m(n), \tag{6.34}$$

and obtain a very good candidate for γ, namely

$$\gamma = \min_{m \leq n \leq q+m-1} \tau_n.$$

The procedure to calculate τ_n, which was only vaguely indicated in [**49**], was clarified in an email discussion with [**50**] and reads as follows.

We know that $d_m(m) = 1$ and $d_m(m+1) = \lambda_m$, see (6.14). Hence we define $\tau_m := \lambda_m$. Let $n \in \{m, \dots, q+m-1\}$ and assume that τ_i for $i < n$ are already defined. So we set

$$\widetilde{d}(n) := \prod_{i=m}^{n-1} \tau_i$$

and note that $d_m(n) \geq \widetilde{d}(n)$. Moreover, we have that

$$d_m(n) \geq d_m(n-g) \prod_{\ell=n-g}^{n-1} \tau_\ell = d_m(n-g) \frac{\widetilde{d}(n)}{\widetilde{d}(n-g)}$$

and hence

$$\frac{\widetilde{d}(n-g)}{\widetilde{d}(n)} \geq \frac{d_m(n-g)}{d_m(n)}. \tag{6.35}$$

Recall (6.15) and note that if we define the set $\mathcal{H}_j(w_i)$ as in (6.27), we can write here also

$$\mathcal{H}^{(w_i)}(n+1) = \sum_{\lfloor \frac{(m+1)}{k} \rfloor < j \leq \lfloor \frac{(n+1)}{k} \rfloor} \mathcal{H}_j^{(w_i)}(n+1).$$

Let $n+1 = ak+b$ where $a, b \in \mathbb{N}$ and $0 \leq b < k$. Note that $\lfloor \frac{(n+1)}{k} \rfloor = a$. To simplify the notation we set $j_0 := \lfloor \frac{(m+1)}{k} \rfloor + 1$. Recall that $g(j) = kj - m$ and $g(j_0) = g_0$.

If $b > 0$ we can follow the reasoning of Case (2) and (6.31) and apply (6.32). In detail we get with $\ell = n - m - 1$

$$\begin{aligned}\sum_{j=j_0}^{a}\sum_{i=1}^{s} x_i \mathcal{H}_j^{(w_i)}(n+1) &= \sum_{j=j_0}^{a}\sum_{r=1}^{s} \eta_r(j)|\mathcal{F}^{(w_i)}(n-kj+m)| \\ &= \sum_{g=g_0}^{ak-m}\sum_{r=1}^{s} \eta_r'(g)|\mathcal{F}^{(w_i)}(n-g)| \\ &= \sum_{g=g_0}^{ak+b-1-m}\sum_{r=1}^{s} \eta_r'(g)|\mathcal{F}^{(w_i)}(n-g)| \\ &= \sum_{g=g_0}^{n-m}\sum_{r=1}^{s} \eta_r'(g)|\mathcal{F}^{(w_i)}(n-g)| \\ &\le \sum_{r=1}^{s} \eta_r''(n-m)|\mathcal{F}_m^{(w_r)}(m)| + \sum_{g=g_0}^{n-m-1} \varrho_g d_m(n-g)\end{aligned}$$

If $b = 0$ we can only follow the arguments of (6.31) and apply (6.32) for periods $j < \frac{n+1}{k}$. For the estimation of the period $\frac{n+1}{k} = a$ we proceed as follows. Let

$$W_a(w_i) := \left[(v[1:m])' \mid v \in V_a^{(w_i)}\right] = [y_1, \ldots, y_r].$$

Obviously,

$$|\mathcal{H}_a^{(w_i)}(n+1)| \le \sum_{t=1}^{r} |\mathcal{F}_m^{(y_t)}(m)| = r.$$

With $f_a^{(t)}(w_i) = |\{w_t \in W_a(w_i)\}|$ and $\theta_t(a) := \sum_{i=1}^{s} x_i f_a^{(t)}(w_i)$ we see that

$$\begin{aligned}\sum_{i=1}^{s} x_i|\mathcal{H}_a^{(w_i)}(n+1)| &\le \sum_{i=1}^{s} x_i \sum_{t=1}^{s} f_a^{(t)}(w_i)|\mathcal{F}_m^{(w_t)}(m)| \\ &= \sum_{t=1}^{s}\sum_{i=1}^{s} x_i f_a^{(t)}(w_i)|\mathcal{F}_m^{(w_t)}(m)| \\ &= \sum_{t=1}^{s} \theta_t(a).\end{aligned}$$

Here, by following (6.31) and applying (6.32) with $\ell = n-m-1$, we infer that

$$\begin{aligned}
&\sum_{j=j_0}^{a}\sum_{i=1}^{s} x_i \mathcal{H}_j^{(w_i)}(n+1)\\
&= \sum_{j=j_0}^{a-1}\sum_{t=1}^{s} \eta_t(j)|\mathcal{F}^{(w_i)}(n-kj+m)| + \sum_{i=1}^{s} x_i \mathcal{H}_a^{(w_i)}(n+1)\\
&= \sum_{g=g_0}^{(a-1)k-m}\sum_{t=1}^{s} \eta_t'(g)|\mathcal{F}^{(w_i)}(n-g)| + \sum_{t=1}^{s}\theta_t(a)\\
&= \sum_{g=g_0}^{n-m}\sum_{t=1}^{s} \eta_t'(g)|\mathcal{F}^{(w_i)}(n-g)| + \sum_{t=1}^{s}\theta_t(a)\\
&\le \sum_{t=1}^{s} \eta_t''(n-m)|\mathcal{F}_m^{(w_t)}(m)| + \sum_{g=g_0}^{n-m-1} \varrho_g d_m(n-g) + \sum_{t=1}^{s}\theta_t(a).
\end{aligned}$$

Note that $\eta'(g) = 0$ for $(a-1)k+1-m \le g \le (a-1)k+k-1-m = n-m$. With

$$\Theta(n+1) := \begin{cases} \sum_{t=1}^{s} \eta_t''(n-m) + \sum_{t=1}^{s}\theta_t(a), & n+1 = ak,\\ \sum_{t=1}^{s} \eta_t''(n-m), & \text{otherwise} \end{cases} \tag{6.36}$$

we can write the general case as follows:

$$\sum_{i=1}^{s} x_i |\mathcal{H}_m^{(w_i)}(n+1)| \le \sum_{g=g_0}^{n-m-1} \varrho_g d_m(n-g) + \Theta(n+1) \tag{6.37}$$

This implies with (6.15) and (6.35) that

$$\begin{aligned}
\frac{d_m(n+1)}{d_m(n)} &\ge \lambda_m - \sum_{g=g_0}^{n-m-1} \varrho_g \frac{d_m(n-g)}{d_m(n)} - \frac{\Theta(n+1)}{d_m(n)}\\
&\ge \lambda_m - \sum_{g=g_0}^{n-m-1} \varrho_g \frac{\widetilde{d}_m(n-g)}{\widetilde{d}_m(n)} - \frac{\Theta(n+1)}{\widetilde{d}_m(n)}\\
&= \lambda_m - \frac{1}{\widetilde{d}_m(n)}\left(\sum_{g=g_0}^{n-m-1} \varrho_g \widetilde{d}(n-g) + \Theta(n+1)\right).
\end{aligned}$$

So we define

$$\tau_n := \lambda_m - \frac{1}{\widetilde{d}(n)}\left(\sum_{g=g_0}^{n-m-1} \varrho_g \widetilde{d}(n-g) + \Theta(n+1)\right). \tag{6.38}$$

If the calculation of τ_n for $n \in \{m, \ldots, q+m-1\}$ reveals that $\tau_n \geq \gamma$ we can infer from $d_m(n+1) \geq \tau_n d_m(n)$ that (6.6) holds.

6.3. Words Avoiding Rational Powers

In this section let $\alpha \in \mathbb{Q}$ such that $\alpha = k + \frac{a}{b} > 1$ with $k, a, b \in \mathbb{N}$ where $\gcd(a, b) = 1$. Moreover, let $\mathcal{F} = \mathcal{F}^{(>\alpha)}(\mathbf{A}_\ell)$, where $\mathbf{A}_\ell$ is an alphabet with $\ell \geq 2$ letters such that $\mathcal{F}$ is an infinite set. Furthermore let $m, p, q \in \mathbb{N}$ such that

$$m > 2, \quad p \geq \tfrac{m}{\alpha - \max(k-1,1)} - 1 \quad \text{and} \quad q \geq \lfloor \alpha p \rfloor + 1 - m.$$

6.3.1. Polynomial. Analogously to Section 6.2.1 the derivation of the lower bound requires the definition of a particular polynomial. Here, the proof of Theorem 6.2 below will motivate its definition.

The shortest forbidden powers will play a major role in what follows. Let y be such a power with $\mathrm{per}(y) = j$. Obviously, y has the structure $y = u^k v$, where $|u| = j$ and $\frac{|y|}{j} > \alpha$. We see that

$$|y| = \lfloor \alpha j \rfloor + 1 \quad \text{and} \quad |v| = \lfloor \tfrac{aj}{b} \rfloor + 1.$$

Let

$$J := \left\{ j \in \mathbb{N} \mid \tfrac{m+1}{\alpha} \leq j \leq \tfrac{q+m}{\alpha} \right\}. \tag{6.39}$$

For every $j \in J$ and $w_i \in \mathcal{F}''(m)$ we define the set of all shortest forbidden powers of minimal period j, which have the word w_i as a suffix and neither contain other forbidden powers as factors nor closed words of length m:

$$\begin{aligned} V_j^{(w_i)} :&= \left\{ y \in \mathcal{L}_m^{(w_i)} \mid y = u^k v \text{ with } u \in \mathcal{F}(j), \mathrm{per}(y) = j \right\} \\ &= \{y_1, \ldots, y_o\}. \end{aligned} \tag{6.40}$$

Now, we define for $1 \leq t \leq o$

$$B_t := \{ b \in \mathbf{A}_\ell \mid b\, y[1 : \lfloor \alpha j \rfloor] \in \mathcal{F}_m \}. \tag{6.41}$$

Note that since $by[1:\lfloor \alpha j \rfloor]$ is α^+-powerfree we know that $B_t \subset \mathbf{A}_\ell \setminus \{y_t[j]\}$. For $j \in J, w_i \in \mathcal{F}''(m)$ and $1 \le t \le o$ we define the multiset

$$U_j(w_i) := \bigcup_{1\le t\le o} U_{j,t}(w_i), \text{ where}$$

$$U_{j,t}(w_i) := [(by_t[1:m-1])' \mid b \in B_t] \subset \mathcal{F}''(m).$$

Let $\eta_r(j) := \sum_{i=1}^{s} x_i e_j^{(r)}(w_i)$, where $e_j^{(r)}(w_i) := |\{w_r \in U_j(w_i)\}|$. We substitute $g(j) := \lfloor \alpha j \rfloor + 1 - m$ and define

$$g(\left\lceil \tfrac{(m+1)}{\alpha} \right\rceil) =: g_0$$

Note that since $\alpha > 1$ the substitution g is injective. Hence, for $g_0 \le g \le q$ we can set

$$\eta_r'(g) := \begin{cases} \eta_r(j), & \text{if there exists a } j \in J \text{ with } g(j) = g \\ 0, & \text{otherwise} \end{cases}.$$

Now, in complete analogy to Section 6.2.1, we define the polynomial

$$\mathcal{P}_m^{(p,q)}(x) := \sum_{g=g_0}^{q} \varrho_g \, x^g, \tag{6.42}$$

where the coefficients are inductively determined as in (6.21), (6.22) and (6.23).

6.3.2. The Lower Bound. The main result of this section is the following Theorem 6.2. It is analogous to Theorem 6.1 and provides a way to calculate lower bounds for the entropy of α^+-powerfree words.

THEOREM 6.2. *Let $\alpha \in \mathbb{Q}$ such that $\alpha = k + \frac{a}{b} > 1$ with $k, a, b \in \mathbb{N}$ where $\gcd(a,b) = 1$. Moreover, let $\mathcal{F} = \mathcal{F}^{(>\alpha)}(\mathbf{A}_\ell)$, where $\mathbf{A}_\ell$ is a finite alphabet with $\ell \ge 2$ letters such that $\mathcal{F}$ is an infinite set. For $m > 2$ let Δ_m be the corresponding matrix from Definition 4.18. Let Δ_m be irreducible, denote its Perron eigenvalue by λ_m, its strictly positive right eigenvector by $(x_1, \ldots, x_s)$, and define $\mu := \frac{\max_{1\le i\le s} x_i}{\min_{1\le i\le s} x_i}$. Furthermore, for $p, q \in \mathbb{N}$ with $p \ge \frac{m}{\alpha - \max(k-1,1)} - 1$ and $q \ge \lfloor \alpha p \rfloor + 1 - m$ construct the polynomial $\mathcal{P}_m^{(p,q)}$*

from (6.42). *If for some* $\gamma \in (1, \lambda_m)$ *and all* $n \in \{m, \dots, q+m-1\}$

$$d_m(n+1) \geq \gamma d_m(n) \tag{6.43}$$

as well as

$$\begin{aligned} &\lambda_m - \mathcal{P}_m^{(p,q)}(\tfrac{1}{\gamma}) - \mu \sum_{j>p} \frac{1}{\gamma^{\lfloor \frac{aj}{b} \rfloor}} \geq \gamma, \quad \text{if} \quad 2 > \alpha > 1 \text{ or} \\ &\lambda_m - \mathcal{P}_m^{(p,q)}(\tfrac{1}{\gamma}) - \frac{1}{\gamma^{(k-1)p-1}(\gamma^{(k-1)}-1)} \geq \gamma, \quad \text{if} \quad \alpha > 2 \end{aligned} \tag{6.44}$$

then

$$h(\mathcal{F}) \geq \log(\gamma).$$

Proof. This theorem is proved almost analogously to Theorem 6.1. Therefore, we restrict ourselves to pointing out what is different in this case.

To show that the inequality (6.43) holds for every $n \in \mathbb{N}$, we argue by induction on $n \geq q+m \geq \lfloor \alpha p \rfloor + 1$. Again, we have to estimate $\sum_{i=1}^{s} x_i |\mathcal{H}^{(w_i)}(n+1)|$ by $d_m(n)$. For $w \in \mathcal{H}^{(w_i)}(n+1)$ denote by

$$y_w = u^k v$$

the shortest forbidden α^+-power, which is a suffix of w. Note that necessarily $|u| = \operatorname{per}(y_w)$. Since

$$\alpha < \tfrac{|y_w|}{\operatorname{per}(y_w)} \leq \tfrac{n+1}{\operatorname{per}(y_w)}$$

we infer that $\operatorname{per}(y_w) < \frac{n+1}{\alpha}$ and hence $\operatorname{per}(y_w) \leq \left\lceil \frac{(n+1)}{\alpha} \right\rceil - 1$.
Because $w[n-m+1 : n+1] \in \mathcal{F}$ and $|w[n-m+1 : n+1]| = m+1$, we know that $|y_w| = \lfloor \alpha \operatorname{per}(y_w) \rfloor + 1 \geq m+2$ and hence $\operatorname{per}(y_w) \geq \left\lceil \frac{(m+1)}{\alpha} \right\rceil$. Overall, we have

$$\left\lceil \tfrac{(m+1)}{\alpha} \right\rceil \leq \operatorname{per}(y_w) \leq \left\lceil \tfrac{(n+1)}{\alpha} \right\rceil - 1. \tag{6.45}$$

In the estimation of $\mathcal{H}_j^{(w_i)}(n+1) := \left\{ w \in \mathcal{H}^{(w_i)}(n+1) \mid \operatorname{per}(y_w) = j \right\}$ we distinguish

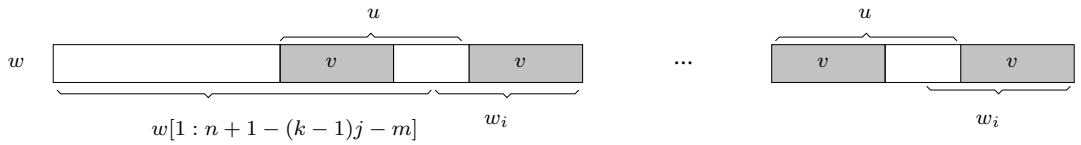


FIGURE 6.4. $k \geq 2, w \in \mathcal{H}_j(w_i)$ where $j \in J_1$, $y = u^k v$ and $|u| = j$.

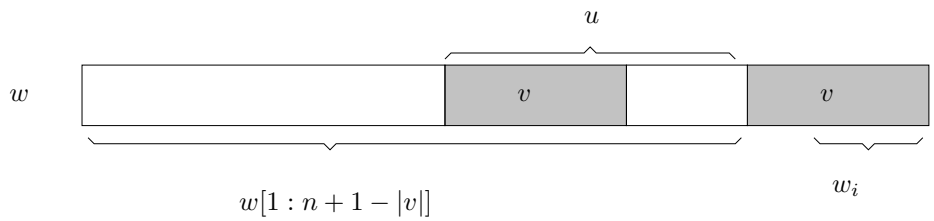


FIGURE 6.5. $k = 1, w \in \mathcal{H}_j(w_i)$ where $j \in J_1$, $y = u^k v$ and $|u| = j$.

(1) $J_1 := \left\{ j \in \mathbb{N} \mid p < j \leq \left\lceil \frac{(n+1)}{\alpha} \right\rceil - 1 \right\}$ and

(2) $J_2 := \left\{ j \in \mathbb{N} \mid \left\lceil \frac{(m+1)}{\alpha} \right\rceil \leq j \leq p \right\}$.

Note that $J_2 \subset J$ from (6.39).

(1) Again consider first the set J_1. Let $j \in J_1$, which means that $(\alpha - 1)j \geq m$. Moreover, let

$$w \in \mathcal{M}_j := \bigcup_{i=1}^{s} \mathcal{H}_j^{(w_i)}(n+1)$$

which is a disjoint union.

If $\alpha > 2$ the word w is completely determined by

$$w[1 : n+1-(|v|+(k-2)|u|+(|u|-|v|)+m)] = w[1 : n+1-(k-1)j-m],$$

compare Figure 6.4. Hence $|\mathcal{H}_j^{(w_i)}(n+1)| \leq |\mathcal{F}_m^{(w_i)}(n+1-(k-1)j)|$, which is exactly (6.29). So we follow the arguments from there, which results in

$$\sum_{i=1}^{s} x_i \Big(\sum_{j \in J_1} |\mathcal{H}_j^{(w_i)}(n+1)| \Big) < \frac{d_m(n)}{\gamma^{(k-1)p-1}(\gamma^{(k-1)} - 1)}.$$

If $\alpha < 2$ we infer that $\alpha = 1 + \frac{a}{b}$ and $\frac{aj}{b} \geq m$. Moreover, w is completely determined by

$$w[1 : n + 1 - |v|] = w[1 : n - \left\lfloor \tfrac{aj}{b} \right\rfloor],$$

see Figure 6.5, and

$$w[n + 2 - m - j : n + 1 - j] = w_i = w[n + 2 - m : n + 1].$$

If we set

$$\mathcal{M}'_j := \left\{ w \in \mathcal{F}_m(n - \left\lfloor \tfrac{aj}{b} \right\rfloor) \mid w[n + 2 - m - j] = 0, w[n + 3 - m - j] = 1 \right\}$$

we clearly see that $|\mathcal{M}_j| \leq |\mathcal{M}'_j|$. Let

$$\mathcal{M}''_j := \left\{ w \in \mathcal{F}_m(n - \left\lfloor \tfrac{aj}{b} \right\rfloor) \mid w[n + 1 - m - \left\lfloor \tfrac{aj}{b} \right\rfloor] = 0, w[n + 2 - m - \left\lfloor \tfrac{aj}{b} \right\rfloor] = 1 \right\}.$$

Obviously, there is a bijection between $\mathcal{M}'_j$ and $\mathcal{M}''_j$, since both fix two letters at different positions. Since $\mathcal{M}''_j = \dot{\bigcup}_{i=1}^{s} \mathcal{F}_m^{(w_i)}(n - \left\lfloor \tfrac{aj}{b} \right\rfloor)$ we infer that

$$|\mathcal{M}_j| \leq |\mathcal{M}'_j| = |\mathcal{M}''_j| = \sum_{i=1}^{s} |\mathcal{F}_m^{(w_i)}(n - \left\lfloor \tfrac{aj}{b} \right\rfloor)|.$$

Moreover,

$$(\min_{1 \leq i \leq s} x_i) \sum_{i=1}^{s} |\mathcal{F}_m^{(w_i)}(n - \left\lfloor \tfrac{aj}{b} \right\rfloor)| \leq \sum_{i=1}^{s} x_i |\mathcal{F}_m^{(w_i)}(n - \left\lfloor \tfrac{aj}{b} \right\rfloor)| = d_m(n - \left\lfloor \tfrac{aj}{b} \right\rfloor)$$

implies that $|\mathcal{M}_j| \leq \dfrac{d_m(n - \left\lfloor \frac{aj}{b} \right\rfloor)}{\min_{1 \leq i \leq s} x_i}$. Overall, we have that

$$\sum_{i=1}^{s} x_i |\mathcal{H}_j^{(w_i)}(n+1)| \leq \max_{1 \leq i \leq s} x_i |\mathcal{M}_j| \leq \mu d_m(n - \left\lfloor \tfrac{aj}{b} \right\rfloor)$$

and hence

$$\sum_{j \in J_1} \sum_{i=1}^{s} x_i |\mathcal{H}_j^{(w_i)}(n+1)| \leq \mu \sum_{j \in J_1} d_m(n - \left\lfloor \tfrac{aj}{b} \right\rfloor) < \mu d_m(n) \sum_{j > p} \frac{1}{\gamma^{\left\lfloor \frac{aj}{b} \right\rfloor}}. \tag{6.46}$$

Case (2) and the rest of the proof is done analogously to the corresponding part of the proof of Theorem 6.1. □

Analogously to the k-powerfree case, we can compute $\tau_n \in \mathbb{R}$ inductively for $n \in \{m, \dots, q+m-1\}$, such that

$$d_m(n+1) \geq \tau_n d_m(n)$$

as follows. Again, we follow the reasoning of the k-powerfree case and concentrate on what is different for α^+-powerfree words.

We know that $d_m(m) = 1$ and $d_m(m+1) = \lambda_m$, see (6.14). Hence we define $\tau_m := \lambda$. Let $n \in \{m, \dots, q+m-1\}$ and assume that τ_i for $i < n$ are already defined. So we set

$$\widetilde{d}(n) := \prod_{i=m}^{n-1} \tau_i.$$

Considering (6.45) we see that

$$\mathcal{H}^{(w_i)}(n+1) = \sum_{\frac{(m+1)}{\alpha} \leq j < \frac{(n+1)}{\alpha}} \mathcal{H}_j^{(w_i)}(n+1).$$

To simplify the notation, we set $j_0 := \left\lceil \frac{(m+1)}{\alpha} \right\rceil$ and $c(n+1) := \left\lceil \frac{n+1}{\alpha} \right\rceil - 1$.

If $\lfloor \alpha c(n+1) \rfloor < n$, we know that $|y| = \lfloor \alpha c(n+1) \rfloor + 1 < n+1$ and hence we can follow the arguments of (6.31) and apply (6.32). Recall that $g(j) = \lfloor \alpha j \rfloor + 1 - m$ and $g(j_0) = g_0$. Note that $g(c) \leq n - m$. In detail, with $\ell = n - m - 1$, we get

$$\begin{aligned}
\sum_{j=j_0}^{c} \sum_{i=1}^{s} x_i \mathcal{H}_j^{(w_i)}(n+1) &= \sum_{j=j_0}^{c} \sum_{r=1}^{s} \eta_r(j) |\mathcal{F}^{(w_i)}(n - \lfloor \alpha j \rfloor - 1 + m)| \\
&= \sum_{g=g_0}^{g(c)} \sum_{r=1}^{s} \eta'_r(g) |\mathcal{F}^{(w_i)}(n-g)| \\
&\leq \sum_{g=g_0}^{n-m} \sum_{r=1}^{s} \eta'_r(g) |\mathcal{F}^{(w_i)}(n-g)| \\
&\leq \sum_{r=1}^{s} \eta''_r(n-m) |\mathcal{F}_m^{(w_r)}(m)| + \sum_{g=g_0}^{n-m-1} \varrho_g d_m(n-g).
\end{aligned}$$

If $\lfloor \alpha c(n+1) \rfloor = n$, we can only follow the reasoning of (6.31) and apply (6.32) for periods $j < c(n+1)$. For the estimation of the period $c := c(n+1)$

we proceed as follows. Define the multiset

$$W_c(w_i) := \left[(v[1:m])' \mid v \in V_c^{(w_i)}\right] = [v_1, \ldots, v_t].$$

Obviously,

$$|\mathcal{H}_c^{(wi)}(n+1)| \le \sum_{r=1}^{t} |\mathcal{F}_m^{(v_r)}(m)| = t.$$

With $f_c^{(r)}(w_i) = |\{w_r \in W_c(w_i)\}|$ and $\theta_r(c) := \sum_{i=1}^{s} x_i f_c^{(r)}(w_i)$ we see that

$$\begin{aligned}\sum_{i=1}^{s} x_i |\mathcal{H}_c^{(w_i)}(n+1)| &\le \sum_{i=1}^{s} x_i \sum_{r=1}^{s} f_c^{(r)}(w_i) |\mathcal{F}_m^{(w_r)}(m)| \\ &= \sum_{r=1}^{s}\sum_{i=1}^{s} x_i f_c^{(r)}(w_i) |\mathcal{F}_m^{(w_r)}(m)| \\ &= \sum_{r=1}^{s} \theta_r(c).\end{aligned}$$

Here, we infer by following the arguments of (6.31) and applying (6.32) with $\ell = n - m - 1$ that

$$\begin{aligned}&\sum_{j=j_0}^{c}\sum_{i=1}^{s} x_i \mathcal{H}_j^{(w_i)}(n+1) \\ &= \sum_{j=j_0}^{c-1}\sum_{r=1}^{s} \eta_r(j) |\mathcal{F}^{(w_i)}(n - \lfloor \alpha j \rfloor - 1 + m)| + \sum_{i=1}^{s} x_i \mathcal{H}_c^{(w_i)}(n+1) \\ &= \sum_{g=g_0}^{g(c-1)}\sum_{r=1}^{s} \eta_r'(g) |\mathcal{F}^{(w_i)}(n-g)| + \sum_{r=1}^{s} \theta_r(c) \\ &\le \sum_{g=g_0}^{n-m}\sum_{r=1}^{s} \eta_r'(g) |\mathcal{F}^{(w_i)}(n-g)| + \sum_{r=1}^{s} \theta_r(c) \\ &\le \sum_{r=1}^{s} \eta_r''(n-m) |\mathcal{F}_m^{(w_r)}(m)| + \sum_{g=g_0}^{n-m-1} \varrho_g d_m(n-g) + \sum_{r=1}^{s} \theta_r(c).\end{aligned}$$

With

$$\Theta(n+1) := \begin{cases} \sum_{r=1}^{s} \eta_r''(n-m) + \sum_{r=1}^{s} \theta_r(c), & \lfloor \alpha c(n+1) \rfloor = n \\ \sum_{r=1}^{s} \eta_r''(n-m), & \text{otherwise} \end{cases} \tag{6.47}$$

we follow exactly the reasoning of the k-powerfree case, which results in

$$\tau_n := \lambda - \frac{1}{\widetilde{d}(n)} \left(\sum_{g=g_0}^{n-m-1} \varrho_g \widetilde{d}(n-g) + \Theta(n+1) \right). \tag{6.48}$$

Again, if the calculation of τ_n for $n \in \{m, \dots, q+m-1\}$ reveals that $\tau_n \geq \gamma$ we can infer from $d_m(n+1) \geq \tau_n d_m(n)$ that (6.6) holds.

In the following chapter, we apply Theorems 6.1 and 6.2 to the six cases introduced above in Section 4.3.1.

CHAPTER 7

Computational Application and Bounds

In this chapter we consider upper and lower bounds for the entropy of the following sets, which were introduced in Section 4.3.1.:

7.1 $\mathcal{F}^{(3)}(\mathbf{A}_2)$ - Cubefree words over $\mathbf{A}_2 = \{0, 1\}$

7.2 $\mathcal{F}^{(2)}(\mathbf{A}_3)$ - Squarefree words over $\mathbf{A}_3 = \{0, 1, 2\}$

7.3 $\mathcal{F}^{(>\frac{7}{4})}(\mathbf{A}_3)$ - $(\frac{7}{4})^+$-powerfree words over $\mathbf{A}_3$

7.4 $\mathcal{F}^{(>\frac{7}{3})}(\mathbf{A}_2)$ - $(\frac{7}{3})^+$-powerfree words over $\mathbf{A}_2$

7.5 $\mathcal{F}^{(2)}(\mathbf{A}_4)$ - Squarefree words over $\mathbf{A}_4 = \{0, 1, 2, 3\}$

7.6 $\mathcal{F}^{(>\frac{7}{5})}(\mathbf{A}_4)$ - $(\frac{7}{5})^+$-powerfree words over $\mathbf{A}_4$

We start with the two classical cases, binary cubefree words in Section 7.1 and ternary squarefree words in Section 7.2. We review the bounds derived by the various approaches mentioned in previous chapters and present the best upper and lower bounds known so far. For the lower bounds our calculations confirm Kolpakov's results from [**49**] exactly. However, for ternary minimally repetitive words, see Section 7.3, our intermediate results differ slightly from Kolpakov's, but nevertheless the resulting lower bound coincides with his. Moreover, in Sections 7.4, 7.5 and 7.6 we analyse sets which have not been studied yet with respect to finding the best upper and lower bounds for their entropy.

The derivation of the bounds relies on computations that were realised in the programming language Python. The code is available on request.

7.1. Binary Cubefree Words

Define for this section $h := h(\mathcal{F}^{(3)}(\mathbf{A}_2))$ as the entropy of cubefree words over the alphabet $\mathbf{A}_2$ and $c(n) := c_{\mathcal{F}^{(3)}(\mathbf{A}_2)}(n)$ as the number of binary cubefree words of length n. The values for $c(n)$ with $n \leq 47$ are given

in [**31**]; an extended list for $n \leq 80$ is shown in Table 7.1. Moreover, the sequence is listed as [**78**, entry A028445].

TABLE 7.1. The number $c(n)$ of binary cubefree words of length n for $n \leq 80$.

n	$c(n)$	n	$c(n)$	n	$c(n)$	n	$c(n)$
1	2	21	7754	41	14565048	61	27286212876
2	4	22	11320	42	21229606	62	39771765144
3	6	23	16502	43	30943516	63	57970429078
4	10	24	24054	44	45102942	64	84496383550
5	16	25	35058	45	65741224	65	123160009324
6	24	26	51144	46	95822908	66	179515213688
7	36	27	74540	47	139669094	67	261657313212
8	56	28	108664	48	203577756	68	381385767316
9	80	29	158372	49	296731624	69	555899236430
10	118	30	230800	50	432509818	70	810266077890
11	174	31	336480	51	630416412	71	1181025420772
12	254	32	490458	52	918879170	72	1721435861086
13	378	33	714856	53	1339338164	73	2509125828902
14	554	34	1041910	54	1952190408	74	3657244826158
15	802	35	1518840	55	2845468908	75	5330716904964
16	1168	36	2213868	56	4147490274	76	7769931925578
17	1716	37	3226896	57	6045283704	77	11325276352154
18	2502	38	4703372	58	8811472958	78	16507465616784
19	3650	39	6855388	59	12843405058	79	24060906866922
20	5324	40	9992596	60	18720255398	80	35070631260904

7.1.1. Upper Bound. According to (4.9), the best upper limit for the entropy h, based on Table 7.1, is

$$h \leq \frac{\log c(80)}{80} \approx 0.389855.$$

For comparison, the limit obtained using the number of words of length 79 is 0.390020, which indicates that these limits are still considerably larger than the actual value of h.

Already in 1983, Brandenburg showed that

$$2 \cdot 2^{\frac{n}{9}} \leq c(n) \leq 2 \cdot 1251^{\frac{n}{17}}$$

which leads in our setting to $0.07702 \leq h \leq 0.41951$. The currently best upper bound is due to Edlin [**31**] and Ochem and Reix [**64**]. They obtained an upper bound on the growth rate of 1.457579201, which corresponds to the bound

$$h \leq 0.376776978 \tag{7.1}$$

on the entropy. We managed to calculate the Perron eigenvalue of the matrix Δ_m from Definition 4.18 up to $m = 40$. According to Theorem 4.19, $\lambda_{40} = 1.457587595$ results in the upper bound

$$h \leq 0.376782736$$

which is only slightly greater than the bound of (7.1). We would expect to improve this value by calculating λ_m for greater m, however this requires huge computational effort.

7.1.2. Lower Bound. We now move on to the lower bound and cubefree morphisms. We have already seen one example, the Thue-Morse morphism, recall (0.1), which is a cubefree morphism from a binary alphabet to a binary alphabet. As explained in Section 4.3.2, it is useful to find uniform cubefree morphisms from larger alphabets, because these provide lower bounds on the entropy. Clearly, if we have an r-uniform cubefree morphism $\varrho\colon \mathbf{A}_\ell^* \to \mathbf{A}_2^*$, it is completely specified by the ℓ words w_i, $1 \leq i \leq \ell$, which are the images of the letters in $\mathbf{A}_\ell$. Since any permutation of the letters in $\mathbf{A}_\ell$ will again yield a uniform cubefree morphism, the set $\{w_1, \ldots, w_\ell\} \subset \mathbf{A}_2^*(r)$ of generating words determines the morphism up to permutation of the letters in $\mathbf{A}_\ell$.

Moreover, the set $\{\overline{w_1}, \ldots, \overline{w_r}\}$, where $\overline{w}$ denotes the image of w under the permutation $0 \leftrightarrow 1$, also defines cubefree morphisms, as does $\{\widetilde{w_1}, \ldots, \widetilde{w_r}\}$, where $\widetilde{w}$ denotes the reversal of w, i.e., the words w read backwards. This is

obvious because the test-sets of Theorem 4.9 are invariant under these operations. Unless the words are palindromic (which means that $\widetilde{w} = w$), the set $\{w_1, \ldots, w_r\}$ thus represents four different morphisms (not taking into account permutation of letters in $\mathbf{A}_\ell$), the fourth obtained by performing both operations, yielding $\{\widetilde{\overline{w_1}}, \ldots, \widetilde{\overline{w_r}}\}$.

For cubefree morphisms from a three-letter alphabet $\mathbf{A}_3$ to two letters, one needs words of length at least six. For length six, there are twelve inequivalent (with respect to the permutation of letters in $\mathbf{A}_3$) cubefree morphisms. The corresponding sets of generating words are

$$\{w_1, w_2, w_4\}, \quad \{w_2, \overline{w_3}, \widetilde{\overline{w_3}}\}, \quad \{w_2, \overline{w_3}, w_4\},$$

and the corresponding images under the two operations explained above. Here, the four words are

$$w_1 = 001011, \quad w_2 = 001101, \quad w_3 = 010110, \quad w_4 = 011001.$$

It turns out that none of these morphisms actually satisfy the sufficient criterion of Theorem 4.7, but cubefreeness was verified using the test set of Theorem 4.9.

One has to go to length nine to find cubefree morphisms from four to two letters. There are 16 in-equivalent morphisms with respect to permutations of the four letters. Explicitly, they are given by the generating sets

$$\begin{aligned} &\{w_1, w_2, \widetilde{\overline{w_2}}, \widetilde{\overline{w_3}}\},\ \{w_4, \overline{w_6}, \overline{w_7}, \overline{w_9}\},\ \{w_5, \overline{w_5}, w_8, \overline{w_8}\}, \\ &\{w_5, \overline{w_5}, \widetilde{w_8}, \widetilde{\overline{w_8}}\},\ \{\overline{w_6}, \widetilde{w_7}, \widetilde{\overline{w_8}}, w_9\} \end{aligned} \tag{7.2}$$

with words

$$\begin{aligned} &w_1 = 001001101, && w_2 = 001010011, && w_3 = 001011001, \\ &w_4 = 001101001, && w_5 = 010010110, && w_6 = 010011010, \\ &w_7 = 010100110, && w_8 = 011001001, && w_9 = 011010110. \end{aligned}$$

Note that $w_9 = \widetilde{w_9}$ is a palindrome, and that two of the five sets are invariant under the permutation $0 \leftrightarrow 1$, which explains why they only represent 16 different morphisms.

Beyond four letters, the test set of Theorem 4.9 becomes unwieldy, but the sufficient criterion of Theorem 4.7 can be used to obtain morphisms. However, these may not have the optimal length, as the examples here show – again for length nine all morphisms violate the conditions of Theorem 4.7. Still, this need not be the case; for instance, morphisms from a five-letter alphabet that satisfy the sufficient criterion exist for length 12, which in this case is the optimal length.

As a consequence of Theorem 4.13, the morphisms (7.2) from a four letter alphabet show that the entropy of cubefree binary words is positive, and that

$$h \geq \frac{\log 2}{8} \approx 0.08664.$$

Using the sufficient condition, this bound can be improved. For instance, for length 15, one can find cubefree morphisms from 10 letters, which yields a lower bound of

$$h \geq \frac{\log 5}{14} \approx 0.11496.$$

However, compared to the best upper bound in (7.1) this bound is unsatisfactory.

Much better lower bounds are delivered by Theorem 6.1. For $m = 35$, we calculated the matrix Δ_m, whose dimension is $|\mathcal{F}''(35)| = 732274$, and checked computationally that it is irreducible and aperiodic. Its Perron eigenvalue is $\lambda_{35} = 1.45759871346$. Since Theorem 6.1 requires that $p \geq m = 35$ and $q \geq 3p - m$ we choose $p = 35$ and $q = 70$. We calculated $\mathcal{P}_m^{(p,q)}$ of (6.20) with the following result.

$$\begin{aligned}\mathcal{P}_{35}^{(35,70)}(x) = {} & 0.890340372\,x^{35} + 1.398381650\,x^{37} + 1.096456371\,x^{38} + \\ & 30.292784422\,x^{40} + 2.533686573\,x^{41} + 1.296919173\,x^{42} + \\ & 28.893958329\,x^{43} + 22.780261551\,x^{44} + 10.699704398\,x^{45} + \\ & 64.314464491\,x^{47} + 92.853910037\,x^{49} + 91.743094160\,x^{50} + \\ & 67.688386637\,x^{51} + 48.613344815\,x^{52} + 68.285930293\,x^{53} + \\ & 113.239315726\,x^{54} + 144.612325329\,x^{56} + 346.136318272\,x^{58} + \\ & 173.468149479\,x^{59} + 465.000387913\,x^{60} + 134.993652948\,x^{61} + \\ & 224.831968847\,x^{62} + 585.928350644\,x^{63} + 355.591901030\,x^{65} + \\ & 1335.518621400\,x^{67} + 343.074472809\,x^{68} + 2202.468158894\,x^{69} + \\ & 11098.126368881\,x^{70}\end{aligned}$$

This polynomial coincides with Kolpakov's polynomial in [**49**].

For the results of the calculation of τ_n for $n \in \{35, \ldots, 104\}$ see Table 7.2. As $\tau_{38} = 1.457567648$ is the lowest value we conclude with $\tau_{38} > \gamma = 1.457567$ that

$$d_{35}(n+1) \geq \gamma d_{35}(n)$$

for all $n \in \{35, \ldots, 104\}$. It is easy to check computationally that

$$\lambda_{35} - \mathcal{P}_{35}^{(35,70)}\left(\tfrac{1}{\gamma}\right) - \frac{1}{\gamma^{69}(\gamma^2-1)} \geq \gamma$$

and hence Theorem 6.1 gives the lower bound

$$\log(\gamma) \approx 0.376768607,$$

which is the best lower bound so far. The difference between this bound and the best upper bound, see (7.1), is less than 8.4×10^{-6}, showing the huge improvement over the previously available estimates.

TABLE 7.2. τ_n for binary cubefree words of length $n \in \{35, \ldots, 104\}$, τ_{38} has the lowest value in the table.

n	τ_n	n	τ_n	n	τ_n
35	1.457598713	59	1.457577241	83	1.457577282
36	1.457598713	60	1.457577338	84	1.457577282
37	1.457598713	61	1.457577353	85	1.457577282
38	**1.457567648**	62	1.457577244	86	1.457577282
39	1.457598713	63	1.457577306	87	1.457577282
40	1.457586935	64	1.457577306	88	1.457577282
41	1.457571424	65	1.457577274	89	1.457577282
42	1.457581866	66	1.457577300	90	1.457577282
43	1.457582750	67	1.457577295	91	1.457577282
44	1.457572786	68	1.457577280	92	1.457577282
45	1.457579086	69	1.457577290	93	1.457577282
46	1.457578947	70	1.457577292	94	1.457577282
47	1.457576440	71	1.457577281	95	1.457577282
48	1.457578364	72	1.457577287	96	1.457577282
49	1.457578158	73	1.457577286	97	1.457577282
50	1.457576963	74	1.457577283	98	1.457577282
51	1.457577784	75	1.457577285	99	1.457577282
52	1.457577857	76	1.457577284	100	1.457577282
53	1.457576839	77	1.457577283	101	1.457577282
54	1.457577482	78	1.457577283	102	1.457577282
55	1.457577485	79	1.457577284	103	1.457577282
56	1.457577170	80	1.457577282	104	1.457577282
57	1.457577405	81	1.457577283		
58	1.457577378	82	1.457577283		

COROLLARY 7.1. *The entropy of binary cubefree words is*

$$h(\mathcal{F}^{(3)}(\mathbf{A}_2)) = 0.37677\,(1).$$

7.2. Ternary Squarefree Words

For this section, let the entropy of squarefree words over the alphabet $\mathbf{A}_3$ be denoted by $h := h(\mathcal{F}^{(2)}(\mathbf{A}_3))$ and the number of ternary squarefree words of length n by $c(n) := c_{\mathcal{F}^{(2)}(\mathbf{A}_3)}(n)$. See [**4**] for a list of $c(n)$ for $n \leq 90$ and [**36**] for a list of $c(n)$ for $91 \leq n \leq 110$. Moreover, the sequence is listed as [**78**, entry A06156].

7.2.1. Upper Bounds. Already in 1983, Brandenburg [**18**] showed that

$$6 \cdot 2^{\frac{n}{22}} \leq a(n) \leq 6 \cdot 1172^{\frac{n}{22}}$$

which leads in our setting to

$$0.03151 \leq h \leq 0.32120.$$

In 1999, Noonan and Zeilberger [**63**] lowered the upper bound to 0.26391 by means of generating functions for the number of words avoiding squares of up to length 23. Grimm and Richard [**75**] used the same method to improve the upper bound to 0.263855. At the moment, the best known upper bound is

$$h \leq 0.263740 \tag{7.3}$$

which was established by Ochem in 2006 using an approach based on the transfer matrix method, see [**64**] for details.

We managed to calculate the Perron eigenvalue of the matrix Δ_m from Definition 4.18 up to $m = 55$. According to Theorem 4.19, $\lambda_{55} = 1.30183467$ results in the upper bound

$$h \leq 0.263775$$

which is only slightly greater than the bound of (7.3). Again, we would expect to improve this value by calculating λ_m for greater m, however this requires huge computational effort.

7.2.2. Lower Bound. In 1998, Zeilberger showed that a Brinkhuis pair of length 18 exists, which by Theorem 4.13 implies that the entropy is bounded by $h \geq \log(2)/17 \approx 0.04077$ [**32**]. By going to larger alphabets, this was subsequently improved to $h \geq \log(65)/40 \approx 0.10436$ by Grimm [**36**] and $h \geq \log(110)/42 \approx 0.11192$ by Sun [**82**].

Again, the recent work of Kolpakov [**49**] has made a large difference to the lower bounds. We recalculated his results independently for $m = 45, p = 52$ and $q = 60$. The matrix Δ_{45} has dimension $|\mathcal{F}''(45)| = 277316$ and is irreducible as well as aperiodic, which we checked computationally. Its Perron eigenvalue is $\lambda_{45} = 1.30201063562$. We calculated the polynomial $\mathcal{P}_m^{(p,q)}$ of (6.20) with the following result, which coincides with Kolpakov's in [**49**],

$$\begin{aligned}\mathcal{P}_{45}^{(52,60)}(x) \;=\; & 3.759478878\,x^{44} + 3.176743177\,x^{45} + 6.048525515\,x^{46} + \\ & 7.120005082\,x^{48} + 14.679230021\,x^{50} + 41.594269716\,x^{52} + \\ & 37.431675327\,x^{55} + 40.471891525\,x^{56} + 32.780085498\,x^{58} + \\ & 5.235192989\,x^{59} + 275.705550875\,x^{60}.\end{aligned}$$

TABLE 7.3. τ_n for ternary squarefree words of length $n \in \{45, \ldots, 104\}$, τ_{47} has the lowest value in the table.

n	τ_n	n	τ_n	n	τ_n
45	1.302010636	65	1.301767180	85	1.301761297
46	1.302010636	66	1.301778798	86	1.301762443
47	**1.301731377**	67	1.301766723	87	1.301761345
48	1.302010636	68	1.301775867	88	1.301762126
49	1.301756981	69	1.301762840	89	1.301761338
50	1.301886615	70	1.301771850	90	1.301761874
51	1.301787341	71	1.301760827	91	1.301761212
52	1.301866606	72	1.301769119	92	1.301761627
53	1.301801609	73	1.301761444	93	1.301761070
54	1.301848713	74	1.301767196	94	1.301761415
55	1.301789005	75	1.301762138	95	1.301760949
56	1.301837273	76	1.301765986	96	1.301761235
57	1.301775049	77	1.301762703	97	1.301760881
58	1.301817493	78	1.301765280	98	1.301761097
59	1.301754764	79	1.301762266	99	1.301760839
60	1.301798867	80	1.301764391	100	1.301760998
61	1.301752252	81	1.301761830	101	1.301760809
62	1.301789332	82	1.301763605	102	1.301760916
63	1.301757832	83	1.301761342	103	1.301760770
64	1.301781712	84	1.301762917	104	1.301760852

For the results of the calculation of τ_n for $n \in \{45, \ldots, 104\}$ see Table 7.3. As $\tau_{47} = 1.301731377$ is the smallest value we conclude with $\tau_{47} > \gamma =$

1.30173 that

$$d_{45}(n+1) \geq \gamma d_{45}(n)$$

for all $n \in \{45, \ldots, 104\}$. It is easy to check computationally that

$$\lambda_{45} - \mathcal{P}_{45}^{(52,60)}\left(\tfrac{1}{\gamma}\right) - \frac{1}{\gamma^{51}(\gamma-1)} \geq \gamma$$

and hence Theorem 6.1 gives the lower bound

$$\log(\gamma) \approx 0.263694,$$

which is the best lower bound so far. The difference between this bound and the best upper bound, see (7.3), is less than 5×10^{-5}, showing the huge improvement over the previously available estimates.

COROLLARY 7.2. *The entropy of ternary squarefree words is*

$$h(\mathcal{F}^{(2)}(\mathbf{A}_3)) = 0.2637\,(1).$$

7.3. Ternary Minimally Repetitive Words

For this section define $h := h(\mathcal{F}^{(>\frac{7}{4})}(\mathbf{A}_3))$ as the entropy of minimally repetitive words over the alphabet $\mathbf{A}_3$ and $c(n) := c_{\mathcal{F}^{(>\frac{7}{4})}(\mathbf{A}_3)}(n)$ as the number of ternary minimally repetitive words of length n. The values for $c(n)$ with $n \leq 73$ are given in Table 7.4.

7.3.1. Upper Bound. According to (4.9), the best upper limit for the entropy h, based on Table 7.4, is

$$h \leq \frac{\log c(73)}{73} \approx 0.264135331.$$

Our best upper limit obtained by calculating the Perron eigenvalue λ_m of the matrix Δ_m from Definition 4.18, is much better. We managed to calculate λ_m for $m \leq 62$ and by Theorem 4.19 $\lambda_{62} = 1.245878780$ results in the upper bound

$$h \leq 0.219841128, \text{ which is the best upper bound so far.} \tag{7.4}$$

TABLE 7.4. The number $c(n)$ of ternary minimally repetitive words of length n for $n \leq 73$.

n	$c(n)$	n	$c(n)$	n	$c(n)$	n	$c(n)$
1	3	21	2388	41	207864	61	16944084
2	6	22	2952	42	259536	62	21108714
3	12	23	3654	43	323088	63	26298210
4	18	24	4596	44	402192	64	32762700
5	30	25	5754	45	501804	65	40805928
6	42	26	7278	46	625152	66	50831748
7	60	27	9144	47	778902	67	63322434
8	78	28	11424	48	971394	68	78880548
9	108	29	14364	49	1210974	69	98254608
10	144	30	18066	50	1508694	70	122399124
11	186	31	22644	51	1880184	71	152473620
12	240	32	28302	52	2343558	72	189929460
13	312	33	35472	53	2919456	73	236599440
14	420	34	44118	54	3638232		
15	528	35	55128	55	4531668		
16	678	36	68688	56	5644764		
17	888	37	85578	57	7031772		
18	1140	38	106740	58	8761464		
19	1464	39	133536	59	10914990		
20	1854	40	166710	60	13598538		

7.3.2. Lower Bound. We now move on to the lower bound. Again, we recalculated Kolpakov's results from [**49**] independently. Here, for the parameters $m = 42, p = 72$ and $q = 85$. The matrix Δ_{42} has dimension $|\mathcal{F}''(42)| = 36141$ and is irreducible as well as aperiodic, which we checked computationally. Its Perron eigenvalue is $\lambda_{42} = 1.247499694$. Note that $p \geq \frac{m}{\alpha-1} - 1 = 55$ and $q \geq \lfloor \alpha p \rfloor + 1 - m = 85$ as required in Theorem 6.2.

We calculated $\mathcal{P}_{42}^{(72,85)}$ of (6.42) with the following result.

$$\begin{aligned}\mathcal{P}_{42}^{(72,85)}(x) =\ & 1.976267794\,x^{42} + 1.148061630\,x^{44} + 3.519576156\,x^{45} + \\ & 1.741045692\,x^{47} + 9.687624120\,x^{49} + 0.126312256\,x^{50} + \\ & 31.479339159\,x^{52} + 12.284335289\,x^{53} + 21.010556529\,x^{54} + \\ & 24.183001280\,x^{56} + 96.529326821\,x^{61} + 129.216325472\,x^{64} + \\ & 256.213309926\,x^{66} + 14.826730989\,x^{67} + 64.163102822\,x^{68} +\end{aligned}$$

$$6.862804848\,x^{69} + 84.819931474\,x^{70} + 2.337609715\,x^{72} +$$
$$175.026144441\,x^{73} + 41.068101819\,x^{74} + 335.714817969\,x^{75} +$$
$$341.576383690\,x^{78} + 329.970328757\,x^{80} + 694.861207725\,x^{81} +$$
$$771.864597673\,x^{82} + 291.777905655\,x^{83} + 583.575596936\,x^{84} +$$
$$10506.605293083\,x^{85}$$

The coefficients of this polynomial coincide exactly with Kolpakov's coefficients in [**49**] for ϱ_n with $n \leq 80$. The remaining coefficients differ from Kolpakov's as follows: ϱ_{81} is greater, ϱ_{82} is greater, ϱ_{83} is lower, ϱ_{84} is greater and ϱ_{85} is lower. We will see below that despite these differences the resulting lower bound coincides.

For the results of the calculation of τ_n for $n \in \{42, \ldots, 126\}$ see Table 7.5. As $\tau_{54} = 1.245344650$ is the lowest value we conclude with $\tau_{54} > \gamma = 1.245$, which is also Kolpakov's value in [**49**], that for all $n \in \{42, \ldots, 126\}$

$$d_{42}(n+1) \geq \gamma d_{42}(n).$$

According to Theorem 6.2 we have to check that

$$\lambda_{42} - \mathcal{P}_{42}^{(72,85)}\left(\tfrac{1}{\gamma}\right) - \mu \sum_{j>72} \gamma^{-\left\lfloor \frac{3j}{4} \right\rfloor} \geq \gamma. \tag{7.5}$$

The following lemma simplifies the computational verification of (7.5).

LEMMA 7.3. *For $j, p \in \mathbb{N}$ and every real $\gamma > 1$ the following identity holds*

$$\sum_{j>p} \gamma^{-\left\lfloor \frac{3j}{4} \right\rfloor} = \gamma^{-\left\lfloor \frac{3p-1}{4} \right\rfloor}(\gamma-1)^{-1} + \gamma^{-3\left\lfloor \frac{p}{4} \right\rfloor}(\gamma^3-1)^{-1}. \tag{7.6}$$

PROOF. Note that $-\left\lfloor \frac{3(p+1)}{4} \right\rfloor$ is the greatest possible negative exponent of the left side in (7.6). We look at j modulo 4 and see that every integer $n \geq \left\lfloor \frac{3(p+1)}{4} \right\rfloor$ occurs as negative exponent in $\sum_{j>p} \gamma^{-\left\lfloor \frac{3j}{4} \right\rfloor}$. Moreover, for

TABLE 7.5. τ_n for ternary minimally repetitive words of length n for $42 \leq n \leq 127$, τ_{54} has the lowest value in the table.

n	τ_n	n	τ_n	n	τ_n
42	1.247499694	71	1.245637587	100	1.245529782
43	1.245471275	72	1.245675063	101	1.245524851
44	1.247499694	73	1.245623748	102	1.245525750
45	1.245698605	74	1.245654950	103	1.245521973
46	1.246641975	75	1.245595098	104	1.245522254
47	1.246354251	76	1.245638459	105	1.245519195
48	1.246647429	77	1.245594230	106	1.245516168
49	1.245938475	78	1.245589941	107	1.245517864
50	1.245990059	79	1.245613773	108	1.245514273
51	1.246329700	80	1.245592883	109	1.245515117
52	1.245792153	81	1.245601716	110	1.245512780
53	1.246201639	82	1.245587756	111	1.245513185
54	**1.245344650**	83	1.245594919	112	1.245511421
55	1.246028938	84	1.245574100	113	1.245510159
56	1.245676653	85	1.245559153	114	1.245510573
57	1.245775390	86	1.245577124	115	1.245508719
58	1.245921383	87	1.245554647	116	1.245508812
59	1.245890045	88	1.245565693	117	1.245507410
60	1.245884759	89	1.245552938	118	1.245507467
61	1.245806093	90	1.245559058	119	1.245506335
62	1.245858586	91	1.245550777	120	1.245505540
63	1.245689934	92	1.245545712	121	1.245505832
64	1.245555975	93	1.245549581	122	1.245504790
65	1.245760518	94	1.245538886	123	1.245504766
66	1.245609325	95	1.245542117	124	1.245504006
67	1.245723805	96	1.245533412	125	1.245503858
68	1.245635385	97	1.245536549	126	1.245503226
69	1.245706126	98	1.245529402		
70	1.245679198	99	1.245526430		

every integer $n \geq \frac{p+1}{4}$ the exponent $-3n$ occurs. In total we have

$$\sum_{j>p} \gamma^{-\left\lfloor \frac{3j}{4} \right\rfloor} = \sum_{n \geq \left\lfloor \frac{3(p+1)}{4} \right\rfloor} \gamma^{-n} + \sum_{n \geq \left\lceil \frac{p+1}{4} \right\rceil} \gamma^{-3n}.$$

We infer with the geometric series that

$$\sum_{n \geq \left\lfloor \frac{3(p+1)}{4} \right\rfloor} \gamma^{-n} = \gamma^{-\left\lfloor \frac{3(p+1)}{4} \right\rfloor} \sum_{n \geq 0} \gamma^{-n} = \gamma^{-\left\lfloor \frac{3(p+1)}{4} \right\rfloor} \gamma(\gamma-1)^{-1} = \gamma^{-\left\lfloor \frac{3p-1}{4} \right\rfloor} (\gamma-1)^{-1}$$

and

$$\begin{aligned}\sum_{n\geq\left\lceil\frac{p+1}{4}\right\rceil}\gamma^{-3n} &= \gamma^{-3\left\lceil\frac{p+1}{4}\right\rceil}\sum_{n\geq 0}\gamma^{-3n}\\ &= \gamma^{-3\left\lceil\frac{p+1}{4}\right\rceil}\gamma^3(\gamma^3-1)^{-1}\\ &= \gamma^{-3\left\lfloor\frac{p}{4}\right\rfloor}(\gamma^3-1)^{-1},\end{aligned}$$

since $-3\left\lceil\frac{p+1}{4}\right\rceil+3=-3(\left\lceil\frac{p+1}{4}\right\rceil-1)=-3(\left\lceil\frac{p-3}{4}\right\rceil)=-3\left\lfloor\frac{p}{4}\right\rfloor$. Thus we have shown that (7.6) holds. □

The previous lemma shows that we have to check that

$$\lambda_m-\mathcal{P}_m^{(p,q)}\left(\tfrac{1}{\gamma}\right)-\gamma^{-\left\lfloor\frac{3p-1}{4}\right\rfloor}(\gamma-1)^{-1}-\gamma^{-3\left\lfloor\frac{p}{4}\right\rfloor}(\gamma^3-1)^{-1}\geq\gamma, \tag{7.7}$$

With our parameters this means

$$\lambda_{42}-\mathcal{P}_{42}^{(72,85)}\left(\tfrac{1}{\gamma}\right)-\gamma^{-53}(\gamma-1)^{-1}-\gamma^{-54}(\gamma^3-1)^{-1}\geq\gamma.$$

This inequality is easily verified computationally and by Theorem 6.2 we conclude that

$$\log(\gamma)\approx 0.219135529\leq h.$$

This is the best lower bound known so far and confirms Kolpakov's lower bound from [**49**], exactly. The difference between this bound and the best upper bound from (7.4) is less than 7.06×10^{-4}.

Corollary 7.4. *The entropy of ternary minimally repetitive words is*

$$h(\mathcal{F}^{(>\frac{7}{4})}(\mathbf{A}_3))=0.219\,(1).$$

7.4. Binary Quasi Minimally Repetitive Words

For this section denote the entropy of $\frac{7}{3}^+$-powerfree words over the alphabet $\mathbf{A}_2$ by $h:=h(\mathcal{F}^{(>\frac{7}{3})}(\mathbf{A}_2))$ and the number of binary $\frac{7}{3}^+$-powerfree words of length n by $c(n):=c_{\mathcal{F}^{(>\frac{7}{3})}(\mathbf{A}_2)}(n)$. The values for $c(n)$ with $n\leq 78$ are given in Table 7.6. Moreover, they are listed as [**78**, entry A082380].

TABLE 7.6. The number $c(n)$ of $\frac{7}{3}^+$-powerfree words of length $n \leq 78$.

n	$c(n)$	n	$c(n)$	n	$c(n)$	n	$c(n)$
1	2	21	774	41	44030	61	2387784
2	4	22	962	42	53808	62	2914544
3	6	23	1178	43	65744	63	3558140
4	10	24	1432	44	80316	64	4343306
5	14	25	1754	45	98052	65	5302072
6	20	26	2160	46	119742	66	6471694
7	30	27	2660	47	146124	67	7899982
8	38	28	3292	48	178488	68	9642654
9	50	29	4016	49	217980	69	11771026
10	64	30	4908	50	266126	70	14368164
11	86	31	5948	51	324890	71	17538942
12	108	32	7278	52	396592	72	21408520
13	136	33	8868	53	484198	73	26132642
14	178	34	10844	54	590970	74	31898812
15	222	35	13278	55	721484	75	38937940
16	276	36	16230	56	880896	76	47529292
17	330	37	19826	57	1075384	77	58018024
18	408	38	24208	58	1312802	78	70818138
19	500	39	29554	59	1602568		
20	618	40	36088	60	1956162		

7.4.1. Upper Bound. According to (4.9), the best upper limit for the entropy h, based on Table 7.6, is

$$h \leq \frac{\log c(78)}{78} \approx 0.231738791.$$

Again, our best upper limit obtained by calculating the Perron eigenvalue λ_m of the matrix Δ_m from Definition 4.18, is much better. We managed to calculate λ_m for $m \leq 66$. By Theorem 4.19 $\lambda_{66} = 1.220699552$ results in the upper bound

$$h \leq 0.199424097, \tag{7.8}$$

which is the best upper bound so far.

7.4.2. Lower Bound. We applied Theorem 6.2 with the parameters $m = 57, p = 55$ and $q = 72$. The matrix Δ_{57} has dimension $|\mathcal{F}''(57)| =$

371870 and is irreducible as well as aperiodic, which we checked computationally. Its Perron eigenvalue is $\lambda_{57} = 1.220842548$. Note that, as required by Theorem 6.2,

$$p \geq \tfrac{m}{\alpha-1} - 1 = 41.75 \quad \text{and} \quad q \geq \lfloor \alpha p \rfloor + 1 - m = 72.$$

We calculated $\mathcal{P}_{57}^{(55,72)}$ of (6.42) with the following result.

$$\mathcal{P}_{57}^{(55,72)}(x) = 2.236195116\, x^{62} + 3.212053840\, x^{64} + 5.371346804\, x^{65} + 7.689687803\, x^{68} + 350.938869329\, x^{72}$$

The results of the calculation of τ_n for $n \in \{57, \ldots, 128\}$ can be found in Table 7.7. As $\tau_{58} = 1.220541650$ is the lowest value we conclude with

TABLE 7.7. τ_n for $\frac{7}{3}^+$-power free words of length $n \in \{57, \ldots, 128\}$, τ_{58} has the lowest value in the table.

n	τ_n	n	τ_n	n	τ_n
57	1.220842548	81	1.220694586	105	1.220699226
58	**1.220541650**	82	1.220706849	106	1.220699824
59	1.220842548	83	1.220707743	107	1.220699309
60	1.220641920	84	1.220699511	108	1.220699659
61	1.220801417	85	1.220705503	109	1.220699105
62	1.220806137	86	1.220698134	110	1.220699529
63	1.220689418	87	1.220703929	111	1.220699507
64	1.220780090	88	1.220697687	112	1.220699192
65	1.220646517	89	1.220702754	113	1.220699372
66	1.220742702	90	1.220702958	114	1.220699126
67	1.220635005	91	1.220699796	115	1.220699331
68	1.220725588	92	1.220702294	116	1.220699016
69	1.220730259	93	1.220699350	117	1.220699263
70	1.220715045	94	1.220701947	118	1.220699214
71	1.220727806	95	1.220697804	119	1.220699052
72	1.220690910	96	1.220701125	120	1.220699166
73	1.220720876	97	1.220700907	121	1.220699013
74	1.220693933	98	1.220699256	122	1.220699117
75	1.220717221	99	1.220700611	123	1.220698977
76	1.220716079	100	1.220699088	124	1.220699060
77	1.220706987	101	1.220700284	125	1.220699060
78	1.220714970	102	1.220698965	126	1.220698947
79	1.220698278	103	1.220699870	127	1.220699012
80	1.220711415	104	1.220700085	128	1.220698932

$\tau_{58} > \gamma = 1.22045$, that

$$d_{57}(n+1) \geq \gamma d_{57}(n)$$

for all $n \in \{57, \ldots, 128\}$. As $\frac{7}{3} > 2$ we have to check that

$$\lambda_{57} - \mathcal{P}_{57}^{(55,72)}\left(\tfrac{1}{\gamma}\right) - \frac{1}{\gamma^{54}(\gamma-1)} \geq \gamma$$

which is easily done computationally. Hence by Theorem 6.2 we have

$$\log(\gamma) \approx 0.199219643 \leq h,$$

which is the best lower bound so far. The difference between this bound and the best upper bound from (7.8) is less than 2.05×10^{-4}.

Corollary 7.5. *The entropy of binary quasi minimally repetitive words is*

$$h\big(\mathcal{F}^{(>\frac{7}{3})}(\mathbf{A}_2)\big) = 0.199\,(1).$$

7.5. Quaternary Squarefree Words

For this section, define the entropy of squarefree words over the alphabet $\mathbf{A}_4$ as $h := h\big(\mathcal{F}^{(2)}(\mathbf{A}_4)\big)$ and the number of quaternary squarefree words of length n as $c(n) := c_{\mathcal{F}^{(2)}(\mathbf{A}_4)}(n)$. The values for $c(n)$ for $n \leq 25$ are listed as [**78**, entry A051041].

7.5.1. Upper Bound. We managed to calculate the Perron eigenvalue λ_m of the matrix Δ_m from Definition 4.18 for $m \leq 17$. By Theorem 4.19 $\lambda_{17} = 2.621592352$ results in the upper bound

$$h \leq 0.963781901, \tag{7.9}$$

which is the best upper bound so far.

7.5.2. Lower Bound. We applied Theorem 6.1 with the parameters $m = 16, p = 17$ and $q = 18$. The matrix Δ_{16} has dimension $|\mathcal{F}''(17)| =$

453998 and is irreducible as well as aperiodic, which we checked computationally. Its Perron eigenvalue is $\lambda_{16} = 2.621725645$. We calculated $\mathcal{P}_{16}^{(17,18)}$ of (6.42) with the following result.

$$\begin{aligned}\mathcal{P}_{16}^{(17,18)}(x) = {} & 2.588218591\,x^{11} + 2.206175304\,x^{12} + 7.8101292412\,x^{13} + \\ & 109.187811438\,x^{15} + 10.426601002\,x^{16} + 268.818977531\,x^{17} + \\ & 1189.465204643\,x^{18}\end{aligned}$$

For the results of the calculation of τ_n for $n \in \{16, \ldots, 33\}$ see Table 7.8.

TABLE 7.8. τ_n for quaternary squarefree words of length n for $16 \leq n \leq 33$, τ_{33} has the lowest value in the table.

n	τ_n	n	τ_n	n	τ_n
16	2.621725645	22	2.621519960	28	2.621508638
17	2.621517232	23	2.621508237	29	2.621507993
18	2.621725645	24	2.621512451	30	2.621508210
19	2.621509285	25	2.621508158	31	2.621507956
20	2.621538400	26	2.621509728	32	2.621508034
21	2.621509069	27	2.621508037	**33**	**2.621507938**

As $\tau_{33} = 2.621507938$ is the lowest value we conclude with $\tau_{33} > \gamma = 2.6214$, that

$$d_{16}(n+1) \geq \gamma d_{16}(n)$$

for all $n \in \{16, \ldots, 33\}$. It is easy to check computationally that

$$\lambda_{16} - \mathcal{P}_{16}^{(17,18)}\left(\tfrac{1}{\gamma}\right) - \frac{1}{\gamma^{16}(\gamma - 1)} \geq \gamma$$

and hence by Theorem 6.1 we have

$$\log(\gamma) \approx 0.963708526 \leq h,$$

which is the best lower bound known so far. The difference between this bound and the best upper bound from (7.9) is less than 7.4×10^{-5}.

COROLLARY 7.6. *The entropy of quaternary squarefree words is*

$$h(\mathcal{F}^{(2)}(\mathbf{A}_4)) = 0.9637\,(1).$$

7.6. Quaternary Minimally Repetitive Words

For this section define $c(n) := c_{\mathcal{F}^{(>\frac{7}{5})}(\mathbf{A}_4)}(n)$ as the number of quaternary minimally repetitive or $\frac{7}{5}^+$-powerfree words of length n and the corresponding entropy as $h := h\big(\mathcal{F}^{(>\frac{7}{5})}(\mathbf{A}_4)\big)$. The values for $c(n)$ for $n \leq 200$ are given in Table 7.9 and Table 7.10.

7.6.1. Upper Bound. According to (4.9), the best upper limit for the entropy h, based on Table 7.10, is

$$h \leq \frac{\log c(200)}{200} \approx 0.100680311.$$

Again, calculating the Perron eigenvalue λ_m of the matrix Δ_m from Definition 4.18, gives a much better upper limit. We managed to calculate λ_m for $m \leq 92$.

By Theorem 4.19 the eigenvalue $\lambda_{92} = 1.072732872$ results in the upper bound

$$h \leq 0.070209477, \tag{7.10}$$

which is the best upper bound so far.

7.6.2. Lower Bound. As in the previous cases we applied Theorem 6.2. We managed to calculate the set of minimally repetitive words up to length $n \leq 200$, so we chose the parameters

$$m = 80, p = 199 \quad \text{and} \quad q = 199. \tag{7.11}$$

Note that p and q are chosen as low as possible under the assumptions of Theorem 6.2, namely that $p \geq \frac{m}{\alpha-1} - 1 = 199$ and $q \geq \lfloor \alpha p \rfloor + 1 - m = 199$. If we chose $m > 80$, this would require $p \geq 202$ and $q \geq 202$, which already exceeds the maximal length of the set of words we calculated.

The matrix Δ_{80} has dimension $|\mathcal{F}''(80)| = 3102$ and is irreducible as well as aperiodic, which we checked computationally. Its Perron eigenvalue is $\lambda_{80} = 1.072732872$.

TABLE 7.9. The number $c(n)$ of minimally repetitive words of length n for $n \leq 160$.

n	$c(n)$	n	$c(n)$	n	$c(n)$	n	$c(n)$
1	4	41	9600	81	166080	121	2644368
2	12	42	10728	82	178200	122	2836224
3	24	43	11568	83	190920	123	3039888
4	48	44	12720	84	204552	124	3252816
5	72	45	13752	85	219960	125	3476760
6	96	46	15000	86	235872	126	3719616
7	120	47	15792	87	253776	127	3979800
8	168	48	16512	88	272616	128	4255488
9	216	49	17808	89	292392	129	4553784
10	288	50	19320	90	313512	130	4878744
11	384	51	20424	91	335808	131	5222736
12	456	52	21960	92	360744	132	5592864
13	504	53	23304	93	386256	133	5985072
14	600	54	24984	94	413808	134	6402216
15	648	55	26880	95	445056	135	6849600
16	696	56	28728	96	477240	136	7328328
17	792	57	31152	97	512136	137	7846296
18	840	58	33312	98	548832	138	8397288
19	960	59	35928	99	587112	139	8991984
20	1128	60	38784	100	629712	140	9633552
21	1224	61	41472	101	674040	141	10309224
22	1416	62	44496	102	723504	142	11030592
23	1512	63	47280	103	773400	143	11800800
24	1704	64	50784	104	828456	144	12629112
25	1920	65	54456	105	888936	145	13516560
26	2136	66	58056	106	951096	146	14466792
27	2448	67	62568	107	1018032	147	15497112
28	2688	68	66792	108	1089048	148	16598664
29	3048	69	71856	109	1167144	149	17775648
30	3216	70	77088	110	1250520	150	19031256
31	3432	71	82560	111	1336560	151	20359968
32	3864	72	88752	112	1434168	152	21778920
33	4248	73	94824	113	1534896	153	23294064
34	4752	74	102144	114	1642464	154	24928776
35	5258	75	109632	115	1758024	155	26680440
36	5808	76	117408	116	1880592	156	28560528
37	6480	77	126240	117	2012280	157	30595440
38	7056	78	135456	118	2152080	158	32765520
39	7848	79	145176	119	2304720	159	35072136
40	8736	80	155064	120	2469528	160	37531728

TABLE 7.10. The number $c(n)$ of minimally repetitive words of length n for $161 \leq n \leq 200$.

n	$c(n)$	n	$c(n)$	n	$c(n)$	n	$c(n)$
161	40156272	171	78841032	181	154630440	191	303286776
162	42944520	172	84342240	182	165459000	192	324549144
163	45917232	173	90220800	183	177065400	193	347302872
164	49123536	174	96536496	184	189490176	194	371578440
165	52579200	175	103337736	185	202755048	195	397464000
166	56272296	176	110551848	186	216857400	196	425011464
167	60236616	177	118241016	187	231877224	197	454442088
168	64457616	178	126438408	188	247920408	198	485911272
169	68936928	179	135187872	189	265113456	199	519643296
170	73724952	180	144567720	190	283529112	200	555879576

We calculated $\mathcal{P}_m^{(p,q)}$ of (6.42) with the following result.

$$\begin{aligned}\mathcal{P}_{80}^{(199,199)}(x) = {} & 0.874226453\,x^{115} + 0.222395316\,x^{119} + 0.755005665\,x^{122} + \\ & 0.988354125\,x^{123} + 1.762888082\,x^{124} + 1.240696480\,x^{125} + \\ & 0.022483470\,x^{128} + 0.359773091\,x^{131} + 1.872686142\,x^{137} + \\ & 1.897878429\,x^{140} + 5.347598893\,x^{145} + 2.984708373\,x^{150} + \\ & 13.02669727\,x^{154} + 4.782742329\,x^{158} + 3.909829464\,x^{159} + \\ & 4.071725989\,x^{164} + 5.742122845\,x^{169} + 4.481000847\,x^{173} + \\ & 14.69270437\,x^{177} + 0.897880184\,x^{178} + 40.68750493\,x^{185} + \\ & 13.122350868\,x^{187} + 1.988684414\,x^{192} + 23.106705854\,x^{195} + \\ & 979.514863208\,x^{199}\end{aligned}$$

For the results of the calculation of τ_n for $n \in \{80, \ldots, 278\}$ see Table 7.11 and Table 7.12. As $\tau_{98} = 1.066057344$ is the lowest value we conclude with

$$\tau_{98} > \gamma = 1.066,$$

TABLE 7.11. τ_n for quaternary minimally repetitive words of length $n \in \{80, \ldots, 278\}$, part 1, τ_{98} has the lowest value.

n	τ_n	n	τ_n	n	τ_n
80	1.072742925	147	1.070457849	214	1.068946370
81	1.072742925	148	1.070392047	215	1.068999879
82	1.072742925	149	1.069759848	216	1.068977949
83	1.072742925	150	1.070378651	217	1.068942669
84	1.072742925	151	1.070263935	218	1.068923349
85	1.072742925	152	1.070058513	219	1.068914312
86	1.072742925	153	1.070298723	220	1.068919061
87	1.072742925	154	1.070064060	221	1.068866031
88	1.072742925	155	1.070260012	222	1.068812301
89	1.072742925	156	1.070168685	223	1.068897692
90	1.072742925	157	1.070221200	224	1.068824241
91	1.072742925	158	1.070245204	225	1.068812248
92	1.072742925	159	1.070094075	226	1.068735837
93	1.072742925	160	1.070248035	227	1.068800145
94	1.072742925	161	1.069097631	228	1.068737434
95	1.072742925	162	1.068740279	229	1.068780789
96	1.072742925	163	1.069548737	230	1.068784665
97	1.072742925	164	1.069947358	231	1.068781382
98	**1.066057344**	165	1.069784620	232	1.068773256
99	1.072109169	166	1.069646853	233	1.068732031
100	1.072151795	167	1.069889426	234	1.068731442
101	1.072191576	168	1.069757701	235	1.068631130
102	1.066282213	169	1.069625229	236	1.068589796
103	1.071573147	170	1.069294669	237	1.068671884
104	1.071409341	171	1.069803453	238	1.068622111
105	1.071498224	172	1.069322663	239	1.068659400
106	1.071581280	173	1.069477646	240	1.068652121
107	1.071601018	174	1.069710367	241	1.068650030
108	1.071531772	175	1.069287875	242	1.068624462
109	1.070352241	176	1.069141270	243	1.068609053
110	1.067793763	177	1.069502888	244	1.068591362
111	1.071076094	178	1.069605733	245	1.068563940
112	1.067784401	179	1.069568979	246	1.068547499
113	1.069856548	180	1.069546602	247	1.068547857
114	1.070857918	181	1.069574676	248	1.068561506
115	1.070920867	182	1.069493431	249	1.068542929
116	1.070819375	183	1.069296795	250	1.068509977
117	1.070826666	184	1.069099263	251	1.068527294
118	1.070874795	185	1.069487326	252	1.068483430
119	1.070968440	186	1.069298375	253	1.068484142
120	1.070954363	187	1.069420561	254	1.068465886

TABLE 7.12. τ_n for quaternary minimally repetitive words of length $n \in \{80, \ldots, 278\}$, part 2

n	τ_n	n	τ_n	n	τ_n
121	1.070819027	188	1.069466188	255	1.068460861
122	1.070857360	189	1.069196936	256	1.068455795
123	1.070163380	190	1.069162257	257	1.068456655
124	1.067964408	191	1.069302413	258	1.068457743
125	1.070626593	192	1.069332250	259	1.068429497
126	1.069291935	193	1.069226088	260	1.068404744
127	1.070454898	194	1.069255620	261	1.068391608
128	1.070526733	195	1.069328709	262	1.068382332
129	1.070474480	196	1.069282722	263	1.068349618
130	1.070470183	197	1.069102413	264	1.068332035
131	1.070448291	198	1.068931344	265	1.068355941
132	1.070521268	199	1.069210579	266	1.068343289
133	1.070505603	200	1.069018083	267	1.068343378
134	1.070493156	201	1.069145128	268	1.068325582
135	1.070477195	202	1.069162991	269	1.068319016
136	1.070521369	203	1.069134190	270	1.068284705
137	1.070492086	204	1.069101512	271	1.068253747
138	1.070443443	205	1.069147657	272	1.068263296
139	1.070393675	206	1.069132900	273	1.068214794
140	1.070414004	207	1.069095858	274	1.068209627
141	1.070454267	208	1.069039514	275	1.068228221
142	1.070490852	209	1.069071536	276	1.068238263
143	1.070449415	210	1.069028415	277	1.068214505
144	1.070435712	211	1.069029086	278	1.068195082
145	1.070271435	212	1.068957724		
146	1.070480124	213	1.069027985		

that $d_{80}(n+1) \geq \gamma d_{80}(n)$ for all $n \in \{80, \ldots, 278\}$. According to Theorem 6.2 we have to check that

$$\lambda_{80} - \mathcal{P}_{80}^{(199,199)}\left(\tfrac{1}{\gamma}\right) - \mu \sum_{j>199} \frac{1}{\gamma^{\lfloor \frac{2j}{5} \rfloor}} \geq \gamma. \tag{7.12}$$

The following lemma simplifies the computational verification of (7.12).

LEMMA 7.7. *For $j, p \in \mathbb{N}$ and every real $\gamma > 1$ the following identity holds*

$$\sum_{j>p} \gamma^{-\lfloor \frac{2j}{5} \rfloor} = 2\gamma^{-\lfloor \frac{2p-3}{5} \rfloor}(\gamma-1)^{-1} + \gamma^{-2\lfloor \frac{p}{5} \rfloor}(\gamma^2-1)^{-1}. \tag{7.13}$$

PROOF. Note that $-\left\lfloor \frac{2(p+1)}{5} \right\rfloor$ is the greatest possible exponent of the left side in (7.13). We look at j modulo 5 and see that every integer $n \geq \left\lfloor \frac{2(p+1)}{5} \right\rfloor$ occurs as negative exponent twice in $\sum_{j>p} \gamma^{-\lfloor \frac{3j}{4} \rfloor}$. Moreover, for every integer $n \geq \frac{p+1}{5}$ the exponent $-2n$ occurs. In total we have

$$\sum_{j>p} \gamma^{-\lfloor \frac{2j}{5} \rfloor} = 2 \sum_{n \geq \lfloor \frac{2(p+1)}{5} \rfloor} \gamma^{-n} + \sum_{n \geq \lceil \frac{p+1}{5} \rceil} \gamma^{-2n}.$$

With the geometric series we deduce that

$$\sum_{n \geq \lfloor \frac{2(p+1)}{5} \rfloor} \gamma^{-n} = \gamma^{-\lfloor \frac{2(p+1)}{5} \rfloor} \sum_{n \geq 0} \gamma^{-n} = \gamma^{-\lfloor \frac{2(p+1)}{5} \rfloor} \gamma(\gamma-1)^{-1} = \gamma^{-\lfloor \frac{2p-3}{5} \rfloor} (\gamma-1)^{-1}$$

and

$$\begin{aligned} \sum_{n \geq \lceil \frac{p+1}{5} \rceil} \gamma^{-2n} &= \gamma^{-2\lceil \frac{p+1}{5} \rceil} \sum_{n \geq 0} \gamma^{-2n} \\ &= \gamma^{-2\lceil \frac{p+1}{5} \rceil} \gamma^2 (\gamma^2-1)^{-1} \\ &= \gamma^{-2\lfloor \frac{p}{5} \rfloor} (\gamma^2-1)^{-1}, \end{aligned}$$

since $-2\left\lceil \frac{p+1}{5} \right\rceil + 2 = -2(\left\lceil \frac{p+1}{5} \right\rceil - 1) = -2(\left\lceil \frac{p-4}{5} \right\rceil) = -2\lfloor \frac{p}{5} \rfloor$. Thus we have shown that (7.13) hold. □

The previous lemma shows that we have to check that

$$\lambda_m - \mathcal{P}_m^{(p,q)}\left(\tfrac{1}{\gamma}\right) - 2\mu\gamma^{-\lfloor \frac{2p-3}{5} \rfloor}(\gamma-1)^{-1} - \mu\gamma^{-2\lfloor \frac{p}{5} \rfloor}(\gamma^2-1)^{-1} \geq \gamma, \tag{7.14}$$

For our parameters, see (7.11), this means

$$\lambda_{80} - \mathcal{P}_{80}^{(199,199)}\left(\tfrac{1}{\gamma}\right) - 2\mu\gamma^{-79}(\gamma-1)^{-1} - \mu\gamma^{-78}(\gamma^2-1)^{-1} \geq \gamma.$$

Computationally it is easy to check that the former inequality with $\gamma = 1.066$ is *not* true. The difference $\lambda_{80} - \gamma = 0.006732871$ but already the term

$$2\mu\gamma^{-\lfloor \frac{2p-3}{5} \rfloor}(\gamma-1)^{-1} + \mu\gamma^{-2\lfloor \frac{p}{5} \rfloor}(\gamma^2-1)^{-1} \tag{7.15}$$

equals 1.517484148. The value of (7.15) is lower for greater p, but greater for lower γ. For $p \geq 412$ its value is lower than $\lambda_{80} - \gamma$. However, this would

require to calculate the set of of minimally repetitive words up to length 412 at least. The problem here is that the involved datasets become really huge. Already, the set of equivalence classes of minimally repetitive words of length 200 has a size of 4.4 GB as text file.

The case of minimally repetitive words over a four letter alphabet reveals the limitation of the procedure based on Theorem 6.2. The upper bound 0.070209477, compare (7.10), for the entropy shows that it is very low compared to the entropies in the other cases. Since a low entropy is equivalent to a low growth rate of the set it seems reasonable that we need much higher parameters for the estimation of the power containing sets $\mathcal{H}^{(w_i)}(n+1)$ from (6.8). Unfortunately, the parameters we need for the procedure to work are so large that they require sets that are beyond our computational scope.

In general, the procedure of Chapter 6 is superior to the methods introduced in Chapter 4, since it estimates the number of elements of certain power containing sets, rather than constructing subsets of powerfree words.

Bibliography

[1] J.-P. Allouche and J. Shallit. *Automatic Sequences*. Cambridge University Press, Cambridge, 2003.

[2] T. M. Apostol. *Introduction to Analytic Number Theory*. Springer, New York, 1998. 5th corrected printing.

[3] M. Baake. Solution of the coincidence problem in dimensions $d \leq 4$. In *The Mathematics of Long-Range Aperiodic Order*, volume 489 of *NATO-ASI C*, pages 9–44. Kluwer, Dordrecht, 1997, arXiv:math.MG/0605222.

[4] M. Baake, V. Elser, and U. Grimm. The entropy of square-free words. *Math. Comput. Modelling*, 26:13–26, 1997, arXiv:math-ph/9809010v1.

[5] M. Baake and U. Grimm. Bravais colourings of planar modules with N-fold symmetry. *Z. Krist.*, 219:72–80, 2004, arXiv:math.CO/0301021.

[6] M. Baake and U. Grimm. Multiple planar coincidences with N-fold symmetry. *Z. Krist.*, 221:571–581, 2006, arXiv:math.MG/0511306.

[7] M. Baake and U. Grimm. *Theory of Aperiodic Order: A Mathematical Invitation*. In preparation.

[8] M. Baake, U. Grimm, M. Heuer, and P. Zeiner. Coincidence rotations of the root lattice A_4. *European J. Combin.*, 29:1808–1819, 2008, arXiv:math.MG/0709.1341v1.

[9] M. Baake, M. Heuer, and R. V. Moody. Similar sublattices of the root lattice A_4. *J. Algebra*, 320(4):1391–1408, 2008, arXiv:math/0702448v2.

[10] M. Baake, P. Kramer, M. Schlottmann, and D. Zeidler. Planar patterns with fivefold symmetry as sections of periodic structures in 4-space. *Intern. J. Mod. Phys. B*, 4:2217–2268, 1990.

[11] M. Baake and R. V. Moody. Similarity submodules and semigroups. In *Quasicrystals and Discrete Geometry*, volume 10 of *Fields Inst. Monogr.*, pages 1–13. Amer. Math. Soc., Providence, RI, 1998.

[12] M. Baake and R. V. Moody. Similarity submodules and root systems in four dimensions. *Canad. J. Math.*, 51:1258–1276, 1999, arXiv:math.MG/9904028.

[13] M. Baake, P. Pleasants, and U. Rehmann. Coincidence site modules in 3-space. *Discrete Comput. Geom.*, 38:111–138, 2007, arXiv:math.MG/0609793.

[14] M. Baake, R. Scharlau, and P. Zeiner. Similar sublattices of planar lattices. *Canad. J. Math.*, arXiv:math.MG/0908.2558v1. In press.

[15] M Baake and P. Zeiner. Coincidences in four dimensions. *Phil. Mag.*, 88:2025–32, 2008, arXiv:math.MG/0712.0363v1.

[16] D. R. Bean, A. Ehrenfeucht, and G. F. McNulty. Avoidable patterns in strings of symbols. *Pacific J. Math.*, 85(2):261–294, 1979.

[17] J. Berstel. Growth of repetition-free words – a review. *Theoret. Comput. Sci.*, 340(2):280–290, 2005.

[18] F.-J. Brandenburg. Uniformly growing k-th power-free homomorphisms. *Theoret. Comput. Sci.*, 23:69–82, 1983.

[19] J. Brinkhuis. Nonrepetitive sequences on three symbols. *Quart. J. Math. Oxford Ser. (2)*, 34:145–149, 1983.

[20] A. Carpi. On Dejean's conjecture over large alphabets. *Theoret. Comput. Sci.*, 385(1-3):137–151, 2007.

[21] J. W. S. Cassels. *An Introduction to the Geometry of Numbers*. Springer, Berlin, 1997. Corrected reprint.

[22] L. Chen, R. V. Moody, and J. Patera. Non-crystallographic root systems. In *Quasicrystals and Discrete Geometry*, volume 10 of *Fields Inst. Monogr.*, pages 135–178. Amer. Math. Soc., Providence, RI, 1998.

[23] J. H. Conway, E. M. Rains, and N. J. A. Sloane. On the existence of similar sublattices. *Canad. J. Math.*, 51:1300–1306, 1999.

[24] J. H. Conway and N. J. A. Sloane. *Sphere Packings, Lattices and Groups*. Springer, New York, third edition, 1999.

[25] M. Crochemore. Sharp characterizations of squarefree morphisms. *Theoret. Comput. Sci.*, 18:221–226, 1982.

[26] J. Currie. There are ternary circular square-free words of length n for $n \geq 18$. *Electron. J. Combin.*, 9(1):Note 10, 7 pp. (electronic), 2002.

[27] J. Currie and N. Rampersad. Dejean's conjecture holds for $n \geq 27$. Preprint, arXiv:math.CO/0901.3188v2.

[28] J. Currie and N. Rampersad. A proof of Dejean's conjecture. Preprint, arXiv:math.CO/0905.1129v3.

[29] J. Currie and N. Rampersad. Dejean's conjecture holds for $n \geq 30$. *Theoret. Comput. Sci.*, 410(30-32):2885–2888, 2009.

[30] F. Dejean. Sur un théorème de Thue. *J. Combinatorial Theory Ser. A*, 13:90–99, 1972.

[31] A. E. Edlin. The number of binary cube-free words of length up to 47 and their numerical analysis. *J. Differ. Equations Appl.*, 5(4-5):353–354, 1999.

[32] S.B. Ekhad and D. Zeilberger. There are more than $2^{n/17}$ n-letter ternary square-free words. *J. Integer Seq.*, 1:Article 98.1.9, 1998.

[33] N. P. Fogg. *Substitutions in Dynamics, Arithmetics and Combinatorics*. Springer-Verlag, Berlin, 2002.

[34] S. Glied. Similarity and coincidence isometries for modules. *Canad. Math. Bull.* In press.

[35] S. Glied and M. Baake. Similarity versus coincidence rotations of lattices. *Z. Krist.*, 223:770–772, 2008, arXiv:math.MG/0808.0109.

[36] U. Grimm. Improved bounds on the number of ternary square-free words. *J. Integer Seq.*, 4(2):Article 01.2.7, 14 pp. (electronic), 2001, arXiv:math.CO/0105245v3.

[37] U. Grimm and M. Heuer. On the entropy and letter frequencies of powerfree words. *Entropy*, 10(4):590–612, 2008, arXiv:math.CO/0811.2119v1.

[38] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, Oxford, sixth edition, 2008.

[39] M. Heuer. Ähnlichkeitsuntergitter des Wurzelgitters A_4, 2006. Diplomarbeit, Universität Bielefeld.

[40] M. Heuer. Similar sublattices and coincidence rotations of the root lattice A_4 and its dual. *Z. Krist.*, 223:817–821, 2008, arXiv:math.MG/0808.1228.

[41] M. Heuer and P. Zeiner. CSLs of the root lattice A_4. *J. Physics: Conf. Series*, 226(012024), 2010. `http://iopscience.iop.org/1742-6596/226/1/012024`.

[42] C. Huck. A note on coincidence isometries of modules in euclidean space. *Z. Krist.*, 224:341–344, 2009, arXiv:math.MG/0811.3551.

[43] J. Karhumäki and J. Shallit. Polynomial versus exponential growth in repetition-free binary words. *J. Combin. Theory Ser. A*, 105:335–347, 2004, arXiv:math.CO/0304095v1.

[44] V. Keräenen. On the k-freeness of morphisms on free monoids. *Lecture Notes in Computer Science*, 247:180–188, 1987.

[45] V. Keränen. On k-repetition free words generated by length uniform morphisms over a binary alphabet. *Lecture Notes in Computer Science*, 194:338–347, 1985.

[46] A. I. Khinchin. *Mathematical Foundations of Information Theory*. Dover Publications Inc., New York, 1957.

[47] M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol. *The Book of Involutions*. American Mathematical Society, Providence, RI, 1998.

[48] M. Koecher and R. Remmert. Hamilton's quaternions. In *H.-D. Ebbinghaus, et al (Eds.), Numbers*, pages 189–220. Springer, New York, 1991.

[49] R. Kolpakov. Efficient lower bounds on the number of repetition-free words. *J. Integer Seq.*, 10(3):Article 07.3.2, 16 pp. (electronic), 2007.

[50] R. Kolpakov. private communication, 2010.

[51] R. Kolpakov, G. Kucherov, and Y. Tarannikov. On repetition-free binary words of minimal density. *Theoret. Comput. Sci.*, 218:161–175, 1999.

[52] S. Lang. *Algebraic number theory*. Springer-Verlag, New York, second edition, 1994.

[53] S. Lang. *Algebra*. Springer-Verlag, New York, third edition, 2002.

[54] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, Cambridge, 1995.

[55] M. Lothaire. *Combinatorics on Words*. Cambridge University Press, Cambridge, 1997.

[56] M. Lothaire. *Algebraic Combinatorics on Words*. Cambridge University Press, Cambridge, 2002.

[57] M. Lothaire. *Applied Combinatorics on Words*. Cambridge University Press, Cambridge, 2005.

[58] M. Mohammad-Noori and J. Currie. Dejean's conjecture and Sturmian words. *European J. Combin.*, 28(3):876–890, 2007.

[59] R. V. Moody and J. Patera. Quasicrystals and icosians. *J. Phys. A*, 26:2829–2853, 1993.

[60] R. V. Moody and A. Weiss. On shelling E_8 quasicrystals. *J. Number Theory*, 47:405–412, 1994.

[61] H. M. Morse. Recurrent geodesics on a surface of negative curvature. *Trans. Amer. Math. Soc.*, 22(1):84–100, 1921.

[62] J. Neukirch. *Algebraic number theory*. Springer-Verlag, Berlin, 1999.

[63] J. Noonan and D. Zeilberger. The Goulden-Jackson cluster method: Extensions, applications, and implementations. *J. Difference Eq. Appl.*, 5:355–377, 1999, arXiv:math.CO/9806036v1.

[64] P. Ochem. A generator of morphisms for infinite words. *Theor. Inform. Appl.*, 40(3):427–441, 2006.

[65] P. Ochem. Letter frequency in infinite repetition-free words. *Theoret. Comput. Sci.*, 380(3):388–392, 2007.

[66] P. Ochem. Unequal letter frequencies in ternary square-free words. *Proceedings of WORDS 2007, Marseille*, 2007. `http://edutice.archives-ouvertes.fr/docs/00/30/71/23/PDF/ochem_cirm.pdf`.

[67] P. Ochem and T. Reix. Upper bound on the number of ternary square-free words. *Workshop on Words and Automata (WOWA'06), St. Petersburg*, 2006. `http://www.lri.fr/perso/~ochem/morphisms/wowa.ps`.

[68] J. M. Ollagnier. Proof of Dejean's conjecture for alphabets with 5, 6, 7, 8, 9, 10 and 11 letters. *Theor. Comput. Sci.*, 95(2):187–205, 1992.

[69] J. Pansiot. A propos d'une conjecture de F. Dejean sur les répétitions dans les mots. *Discrete Appl. Math.*, 7(3):297–311, 1984.

[70] K. Petersen. *Ergodic Theory*. Cambridge University Press, Cambridge, 1989. Corrected reprint.

[71] P. A. B. Pleasants, M. Baake, and J. Roth. Planar coincidences for N-fold symmetry. *J. Math. Phys.*, 37:1029–1058, 1996, arXiv:math.MG/0511147.

[72] M. Queffélec. *Substitution Dynamical Systems—spectral analysis*. Springer-Verlag, Berlin, 1987.

[73] M Rao. Last cases of Dejean's conjecture. *Proceedings of WORDS 2009, Salerno*, 2009. `http://www.labri.fr/perso/rao/publi.php?lang=fr`.

[74] I. Reiner. *Maximal Orders*. Clarendon Press, Oxford, 2003. Corrected reprint.

[75] C. Richard and U. Grimm. On the entropy and letter frequencies of ternary square-free words. *Electron. J. Combin.*, 11(1):Research Paper 14, 19 pp. (electronic), 2004, arXiv:math.CO/0302302v2.

[76] G. Richomme and F. Wlazinski. Some results on k-power-free morphisms. *Theoret. Comput. Sci.*, 273:119–142, 2002.

[77] G. Richomme and F. Wlazinski. Existence of finite test-sets for k-powerfreeness of uniform morphims. *Discrete Applied Math.*, 155:2001–2016, 2007, arXiv:cs.DM/0512051v1.

[78] N. J. A. Sloane. The online encyclopedia of integer sequences. `http://www.research.att.com/~njas/sequences/`.

[79] U. Staemmler. Idealklassenzahlen von Quaternionenalgebren über algebraischen Zahlkörpern, 2002. Diplomarbeit, Universität des Saarlandes.

[80] R. P. Stanley. *Enumerative Combinatorics Vol. 1*. Cambridge University Press, Cambridge, 1997.

[81] J. M. Steele. *Probability Theory and Combinatorial Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[82] X. Sun. New lower-bound on the number of ternary square-free words. *J Integer Seq.*, 6:Article 03.3.2, 2003.

[83] A. Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. In *Selected Mathematical Papers of Axel Thue*, pages 413–477. Universitetsforlaget, Oslo, 1977. Reprint of 1912 original.

[84] A. Thue. Über unendliche Zeichenreihen. In *Selected Mathematical Papers of Axel Thue*, pages 139–158. Universitetsforlaget, Oslo, 1977. Reprint of 1906 original.

[85] M.-F. Vignéras. *Arithmétique Des Algèbres de Quaternions*, volume 800 of *Lecture Notes in Mathematics*. Springer, Berlin, 1980.

[86] P. Walters. *An Introduction to Ergodic Theory*. Springer-Verlag, New York, 1982.

[87] D. B. Zagier. *Zetafunktionen und Quadratische Körper*. Springer-Verlag, Berlin, 1981.

[88] P. Zeiner. Coincidences of hypercubic lattices in 4 dimensions. *Z. Krist.*, 221:105–114, 2006, arXiv:math.MG/0605526.

[89] P. Zeiner. Multiple CSLs for the body centered cubic lattice. *J. Physics: Conf. Series*, 30:163–167, 2006, arXiv:math.MG/0605521.

[90] P. Zeiner. Multiplicativity in the theory of coincidence site lattices. *J. Physics: Conf. Series*, 226(012025), 2010. `http://iopscience.iop.org/1742-6596/226/1/012025/`.

[91] Y. M. Zou. Indices of coincidence isometries of the hypercubic lattice $\mathbb{Z}^n$. *Acta Crystallogr. Sect. A*, 62:454–458, 2006.

[92] Y. M. Zou. Structures of coincidence symmetry groups. *Acta Crystallogr. Sect. A*, 62:109–114, 2006.

Index